

# Code\_All

Rizal Mujahiddan

5/24/2022

## Analysis Data with R

### Import Library

```
current_0 <- Sys.time()
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.7      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(dplyr)
library(rpart)
library(rpart.plot)
library(caret)

## Loading required package: lattice

##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
##     lift

library(caTools)
library(earth)

## Loading required package: Formula
```

```

## Loading required package: plotmo

## Loading required package: plotrix

## Loading required package: TeachingDemos

library(mda)

## Loading required package: class

## Loaded mda 0.5-3

library(ROSE)

## Loaded ROSE 0.0-4

library(DataExplorer)
library(car)

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##      recode

## The following object is masked from 'package:purrr':
##
##      some

library(randomForest)

## randomForest 4.7-1.1

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:dplyr':
##
##      combine

## The following object is masked from 'package:ggplot2':
##
##      margin

```

```
library(mlr)
```

```
## Loading required package: ParamHelpers
```

```
## Warning message: 'mlr' is in 'maintenance-only' mode since July 2019.  
## Future development will only happen in 'mlr3'  
## (<https://mlr3.ml-org.com>). Due to the focus on 'mlr3' there might be  
## uncaught bugs meanwhile in {mlr} - please consider switching.
```

```
##
```

```
## Attaching package: 'mlr'
```

```
## The following object is masked from 'package:caret':
```

```
##
```

```
##      train
```

## Import data from csv

```
data <- read.csv("dataset.csv")  
ukuran_data <- dim(data)  
head_data <- head(data)  
write.csv(head_data, "head_data.csv")  
summary(data)
```

```
##      encounter_id      patient_id      hospital_id      age  
## Min.      :      1      Min.      :      1      Min.      :  2.0      Min.      :16.00  
## 1st Qu.: 32852      1st Qu.: 32830      1st Qu.: 47.0      1st Qu.:52.00  
## Median : 65665      Median : 65413      Median :109.0      Median :65.00  
## Mean   : 65606      Mean   : 65537      Mean   :105.7      Mean   :62.31  
## 3rd Qu.: 98342      3rd Qu.: 98298      3rd Qu.:161.0      3rd Qu.:75.00  
## Max.   :131051      Max.   :131051      Max.   :204.0      Max.   :89.00  
##                                     NA's      :4228  
##      bmi      elective_surgery      ethnicity      gender  
## Min.      :14.85      Min.      :0.0000      Length:91713      Length:91713  
## 1st Qu.:23.64      1st Qu.:0.0000      Class :character      Class :character  
## Median :27.66      Median :0.0000      Mode  :character      Mode  :character  
## Mean   :29.19      Mean   :0.1837  
## 3rd Qu.:32.93      3rd Qu.:0.0000  
## Max.   :67.81      Max.   :1.0000  
## NA's      :3429  
##      height      icu_admit_source      icu_id      icu_stay_type  
## Min.      :137.2      Length:91713      Min.      : 82.0      Length:91713  
## 1st Qu.:162.5      Class :character      1st Qu.:369.0      Class :character  
## Median :170.1      Mode  :character      Median :504.0      Mode  :character  
## Mean   :169.6      Mean   :508.4  
## 3rd Qu.:177.8      3rd Qu.:679.0  
## Max.   :195.6      Max.   :927.0  
## NA's      :1334  
##      icu_type      pre_icu_los_days      weight      apache_2_diagnosis  
## Length:91713      Min.      :-24.94722      Min.      : 38.60      Min.      :101.0
```

```

## Class :character    1st Qu.: 0.03542    1st Qu.: 66.80    1st Qu.:113.0
## Mode :character    Median : 0.13889    Median : 80.30    Median :122.0
##                               Mean : 0.83577    Mean : 84.03    Mean :185.4
##                               3rd Qu.: 0.40903    3rd Qu.: 97.10    3rd Qu.:301.0
##                               Max. :159.09097    Max. :186.00    Max. :308.0
##                               NA's :2720    NA's :1662
## apache_3j_diagnosis apache_post_operative arf_apache gcs_eyes_apache
## Min. : 0.01    Min. :0.0000    Min. :0.000    Min. :1.000
## 1st Qu.: 203.01    1st Qu.:0.0000    1st Qu.:0.000    1st Qu.:3.000
## Median : 409.02    Median :0.0000    Median :0.000    Median :4.000
## Mean : 558.22    Mean :0.2011    Mean :0.028    Mean :3.465
## 3rd Qu.: 703.03    3rd Qu.:0.0000    3rd Qu.:0.000    3rd Qu.:4.000
## Max. :2201.05    Max. :1.0000    Max. :1.000    Max. :4.000
## NA's :1101    NA's :715    NA's :1901
## gcs_motor_apache gcs_unable_apache gcs_verbal_apache heart_rate_apache
## Min. :1.000    Min. :0.0000    Min. :1.000    Min. : 30.00
## 1st Qu.:6.000    1st Qu.:0.0000    1st Qu.:4.000    1st Qu.: 86.00
## Median :6.000    Median :0.0000    Median :5.000    Median :104.00
## Mean :5.471    Mean :0.0095    Mean :3.995    Mean : 99.71
## 3rd Qu.:6.000    3rd Qu.:0.0000    3rd Qu.:5.000    3rd Qu.:120.00
## Max. :6.000    Max. :1.0000    Max. :5.000    Max. :178.00
## NA's :1901    NA's :1037    NA's :1901    NA's :878
## intubated_apache map_apache resprate_apache temp_apache
## Min. :0.0000    Min. : 40.00    Min. : 4.00    Min. :32.10
## 1st Qu.:0.0000    1st Qu.: 54.00    1st Qu.:11.00    1st Qu.:36.20
## Median :0.0000    Median : 67.00    Median :28.00    Median :36.50
## Mean :0.1512    Mean : 88.02    Mean :25.81    Mean :36.41
## 3rd Qu.:0.0000    3rd Qu.:125.00    3rd Qu.:36.00    3rd Qu.:36.70
## Max. :1.0000    Max. :200.00    Max. :60.00    Max. :39.70
## NA's :715    NA's :994    NA's :1234    NA's :4108
## ventilated_apache d1_diasbp_max d1_diasbp_min d1_diasbp_noninvasive_max
## Min. :0.0000    Min. : 46.00    Min. :13.00    Min. : 46.00
## 1st Qu.:0.0000    1st Qu.: 75.00    1st Qu.:42.00    1st Qu.: 75.00
## Median :0.0000    Median : 86.00    Median :50.00    Median : 87.00
## Mean :0.3257    Mean : 88.49    Mean :50.16    Mean : 88.61
## 3rd Qu.:1.0000    3rd Qu.: 99.00    3rd Qu.:58.00    3rd Qu.: 99.00
## Max. :1.0000    Max. :165.00    Max. :90.00    Max. :165.00
## NA's :715    NA's :165    NA's :165    NA's :1040
## d1_diasbp_noninvasive_min d1_heartrate_max d1_heartrate_min d1_mbp_max
## Min. :13.00    Min. : 58    Min. : 0.00    Min. : 60.0
## 1st Qu.:42.00    1st Qu.: 87    1st Qu.: 60.00    1st Qu.: 90.0
## Median :50.00    Median :101    Median : 69.00    Median :102.0
## Mean :50.24    Mean :103    Mean : 70.32    Mean :104.7
## 3rd Qu.:58.00    3rd Qu.:116    3rd Qu.: 81.00    3rd Qu.:116.0
## Max. :90.00    Max. :177    Max. :175.00    Max. :184.0
## NA's :1040    NA's :145    NA's :145    NA's :220
## d1_mbp_min d1_mbp_noninvasive_max d1_mbp_noninvasive_min d1_resprate_max
## Min. : 22.00    Min. : 60.0    Min. : 22.00    Min. :14.00
## 1st Qu.: 55.00    1st Qu.: 90.0    1st Qu.: 55.00    1st Qu.:22.00
## Median : 64.00    Median :102.0    Median : 64.00    Median :26.00
## Mean : 64.87    Mean :104.6    Mean : 64.94    Mean :28.88
## 3rd Qu.: 75.00    3rd Qu.:116.0    3rd Qu.: 75.00    3rd Qu.:32.00
## Max. :112.00    Max. :181.0    Max. :112.00    Max. :92.00
## NA's :220    NA's :1479    NA's :1479    NA's :385

```

##	d1_resprate_min	d1_spo2_max	d1_spo2_min	d1_sysbp_max
##	Min. : 0.00	Min. : 0.00	Min. : 0.00	Min. : 90.0
##	1st Qu.: 10.00	1st Qu.: 99.00	1st Qu.: 89.00	1st Qu.:130.0
##	Median : 13.00	Median :100.00	Median : 92.00	Median :146.0
##	Mean : 12.85	Mean : 99.24	Mean : 90.45	Mean :148.3
##	3rd Qu.: 16.00	3rd Qu.:100.00	3rd Qu.: 95.00	3rd Qu.:164.0
##	Max. :100.00	Max. :100.00	Max. :100.00	Max. :232.0
##	NA's :385	NA's :333	NA's :333	NA's :159
##	d1_sysbp_min	d1_sysbp_noninvasive_max	d1_sysbp_noninvasive_min	
##	Min. : 41.00	Min. : 90.0	Min. : 41.03	
##	1st Qu.: 83.00	1st Qu.:130.0	1st Qu.: 84.00	
##	Median : 96.00	Median :146.0	Median : 96.00	
##	Mean : 96.92	Mean :148.2	Mean : 96.99	
##	3rd Qu.:110.00	3rd Qu.:164.0	3rd Qu.:110.00	
##	Max. :160.00	Max. :232.0	Max. :160.00	
##	NA's :159	NA's :1027	NA's :1027	
##	d1_temp_max	d1_temp_min	h1_diasbp_max	h1_diasbp_min
##	Min. :35.10	Min. :31.89	Min. : 37.00	Min. : 22.00
##	1st Qu.:36.90	1st Qu.:36.10	1st Qu.: 62.00	1st Qu.: 52.00
##	Median :37.11	Median :36.40	Median : 74.00	Median : 62.00
##	Mean :37.28	Mean :36.27	Mean : 75.36	Mean : 62.84
##	3rd Qu.:37.60	3rd Qu.:36.66	3rd Qu.: 86.00	3rd Qu.: 73.00
##	Max. :39.90	Max. :37.80	Max. :143.00	Max. :113.00
##	NA's :2324	NA's :2324	NA's :3619	NA's :3619
##	h1_diasbp_noninvasive_max	h1_diasbp_noninvasive_min	h1_heartrate_max	
##	Min. : 37.00	Min. : 22.00	Min. : 46.00	
##	1st Qu.: 63.00	1st Qu.: 52.00	1st Qu.: 77.00	
##	Median : 74.00	Median : 62.00	Median : 90.00	
##	Mean : 75.81	Mean : 63.27	Mean : 92.23	
##	3rd Qu.: 87.00	3rd Qu.: 74.00	3rd Qu.:106.00	
##	Max. :144.00	Max. :114.00	Max. :164.00	
##	NA's :7350	NA's :7350	NA's :2790	
##	h1_heartrate_min	h1_mbp_max	h1_mbp_min	h1_mbp_noninvasive_max
##	Min. : 36.00	Min. : 49.00	Min. : 32.0	Min. : 49.00
##	1st Qu.: 69.00	1st Qu.: 77.00	1st Qu.: 66.0	1st Qu.: 77.00
##	Median : 82.00	Median : 90.00	Median : 78.0	Median : 90.00
##	Mean : 83.66	Mean : 91.61	Mean : 79.4	Mean : 91.59
##	3rd Qu.: 97.00	3rd Qu.:104.00	3rd Qu.: 92.0	3rd Qu.:104.00
##	Max. :144.00	Max. :165.00	Max. :138.0	Max. :163.00
##	NA's :2790	NA's :4639	NA's :4639	NA's :9084
##	h1_mbp_noninvasive_min	h1_resprate_max	h1_resprate_min	h1_spo2_max
##	Min. : 32.00	Min. :10.00	Min. : 0.00	Min. : 0.00
##	1st Qu.: 66.00	1st Qu.:18.00	1st Qu.: 14.00	1st Qu.: 97.00
##	Median : 79.00	Median :21.00	Median : 16.00	Median : 99.00
##	Mean : 79.71	Mean :22.63	Mean : 17.21	Mean : 98.05
##	3rd Qu.: 92.00	3rd Qu.:26.00	3rd Qu.: 20.00	3rd Qu.:100.00
##	Max. :138.00	Max. :59.00	Max. :189.00	Max. :100.00
##	NA's :9084	NA's :4357	NA's :4357	NA's :4185
##	h1_spo2_min	h1_sysbp_max	h1_sysbp_min	h1_sysbp_noninvasive_max
##	Min. : 0.00	Min. : 75.0	Min. : 53.0	Min. : 75.0
##	1st Qu.: 94.00	1st Qu.:113.0	1st Qu.: 98.0	1st Qu.:113.0
##	Median : 96.00	Median :131.0	Median :115.0	Median :130.0
##	Mean : 95.17	Mean :133.2	Mean :116.4	Mean :133.1
##	3rd Qu.: 99.00	3rd Qu.:150.0	3rd Qu.:134.0	3rd Qu.:150.0

```

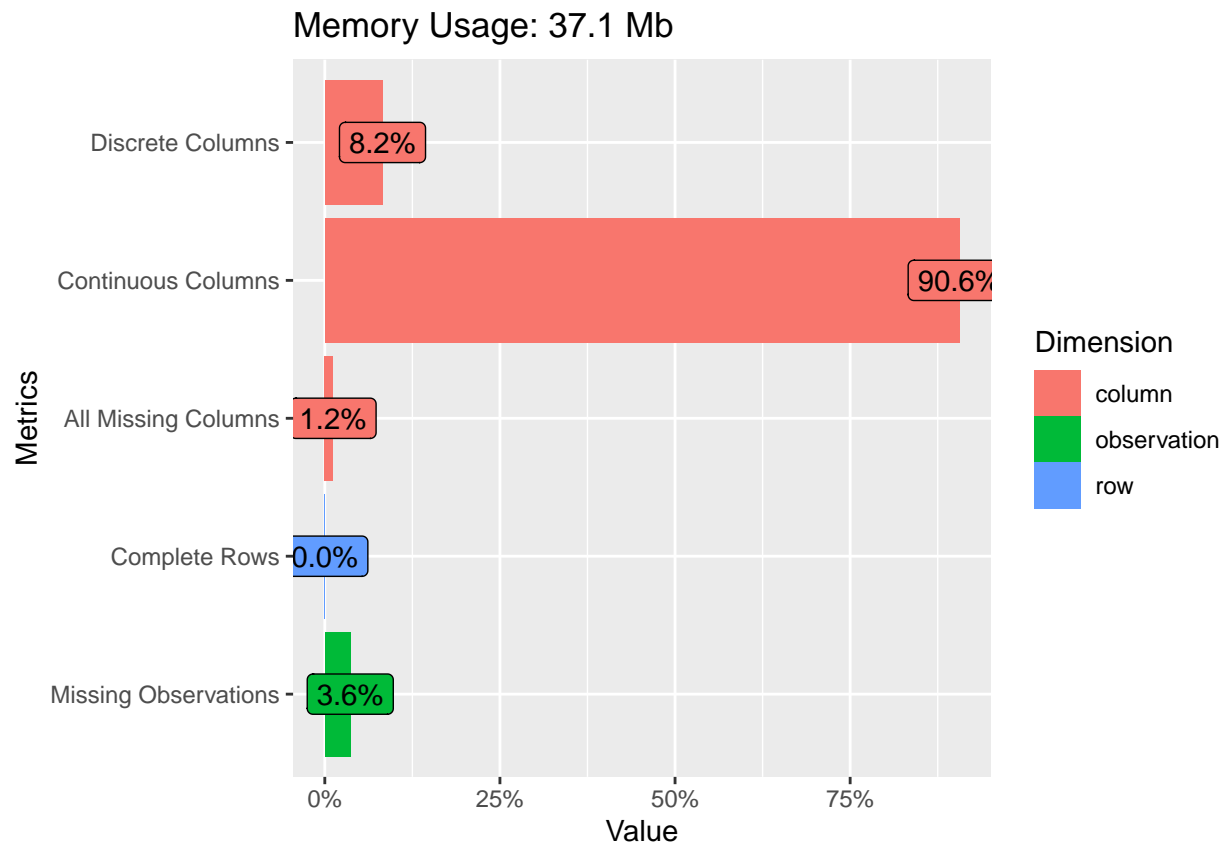
## Max. :100.00 Max. :223.0 Max. :194.0 Max. :223.0
## NA's :4185 NA's :3611 NA's :3611 NA's :7341
## h1_sysbp_noninvasive_min d1_glucose_max d1_glucose_min d1_potassium_max
## Min. : 53.0 Min. : 73.0 Min. : 33.0 Min. :2.800
## 1st Qu.: 98.0 1st Qu.:117.0 1st Qu.: 91.0 1st Qu.:3.800
## Median :115.0 Median :150.0 Median :107.0 Median :4.200
## Mean :116.5 Mean :174.6 Mean :114.4 Mean :4.252
## 3rd Qu.:134.0 3rd Qu.:201.0 3rd Qu.:131.0 3rd Qu.:4.600
## Max. :195.0 Max. :611.0 Max. :288.0 Max. :7.000
## NA's :7341 NA's :5807 NA's :5807 NA's :9585
## d1_potassium_min apache_4a_hospital_death_prob apache_4a_icu_death_prob
## Min. :2.400 Min. :-1.000 Min. :-1.000
## 1st Qu.:3.600 1st Qu.: 0.020 1st Qu.: 0.010
## Median :3.900 Median : 0.050 Median : 0.020
## Mean :3.935 Mean : 0.087 Mean : 0.044
## 3rd Qu.:4.300 3rd Qu.: 0.130 3rd Qu.: 0.060
## Max. :5.800 Max. : 0.990 Max. : 0.970
## NA's :9585 NA's :7947 NA's :7947
## aids cirrhosis diabetes_mellitus hepatic_failure
## Min. :0e+00 Min. :0.0000 Min. :0.0000 Min. :0.000
## 1st Qu.:0e+00 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.000
## Median :0e+00 Median :0.0000 Median :0.0000 Median :0.000
## Mean :9e-04 Mean :0.0157 Mean :0.2252 Mean :0.013
## 3rd Qu.:0e+00 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.000
## Max. :1e+00 Max. :1.0000 Max. :1.0000 Max. :1.000
## NA's :715 NA's :715 NA's :715 NA's :715
## immunosuppression leukemia lymphoma
## Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median :0.0000
## Mean :0.0262 Mean :0.0071 Mean :0.0041
## 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000 Max. :1.0000
## NA's :715 NA's :715 NA's :715
## solid_tumor_with_metastasis apache_3j_bodysystem apache_2_bodysystem
## Min. :0.0000 Length:91713 Length:91713
## 1st Qu.:0.0000 Class :character Class :character
## Median :0.0000 Mode :character Mode :character
## Mean :0.0206
## 3rd Qu.:0.0000
## Max. :1.0000
## NA's :715
## X hospital_death
## Mode:logical Min. :0.0000
## NA's:91713 1st Qu.:0.0000
## Median :0.0000
## Mean :0.0863
## 3rd Qu.:0.0000
## Max. :1.0000
##

```

## Exploring Data

### Mengenai Dataset Itu Sendiri

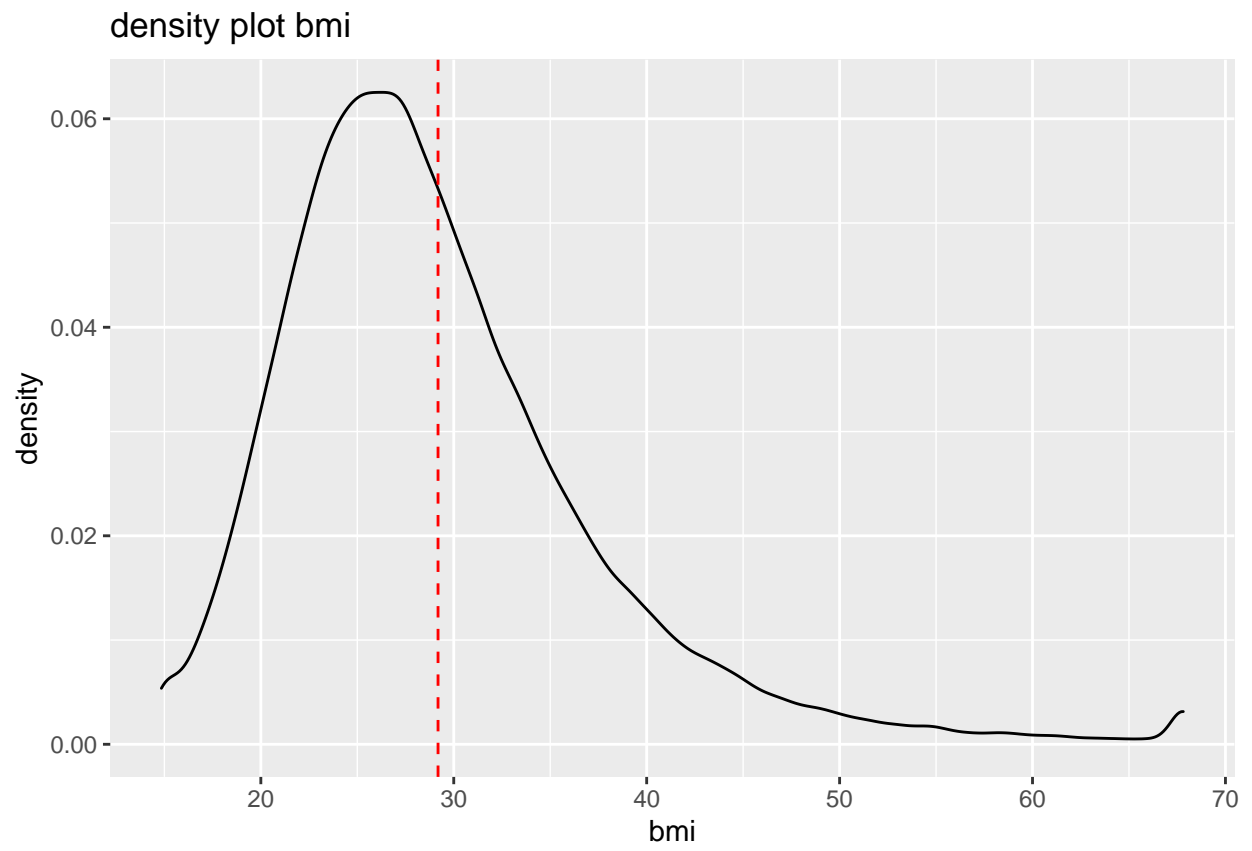
```
DataExplorer::plot_intro(data)
```



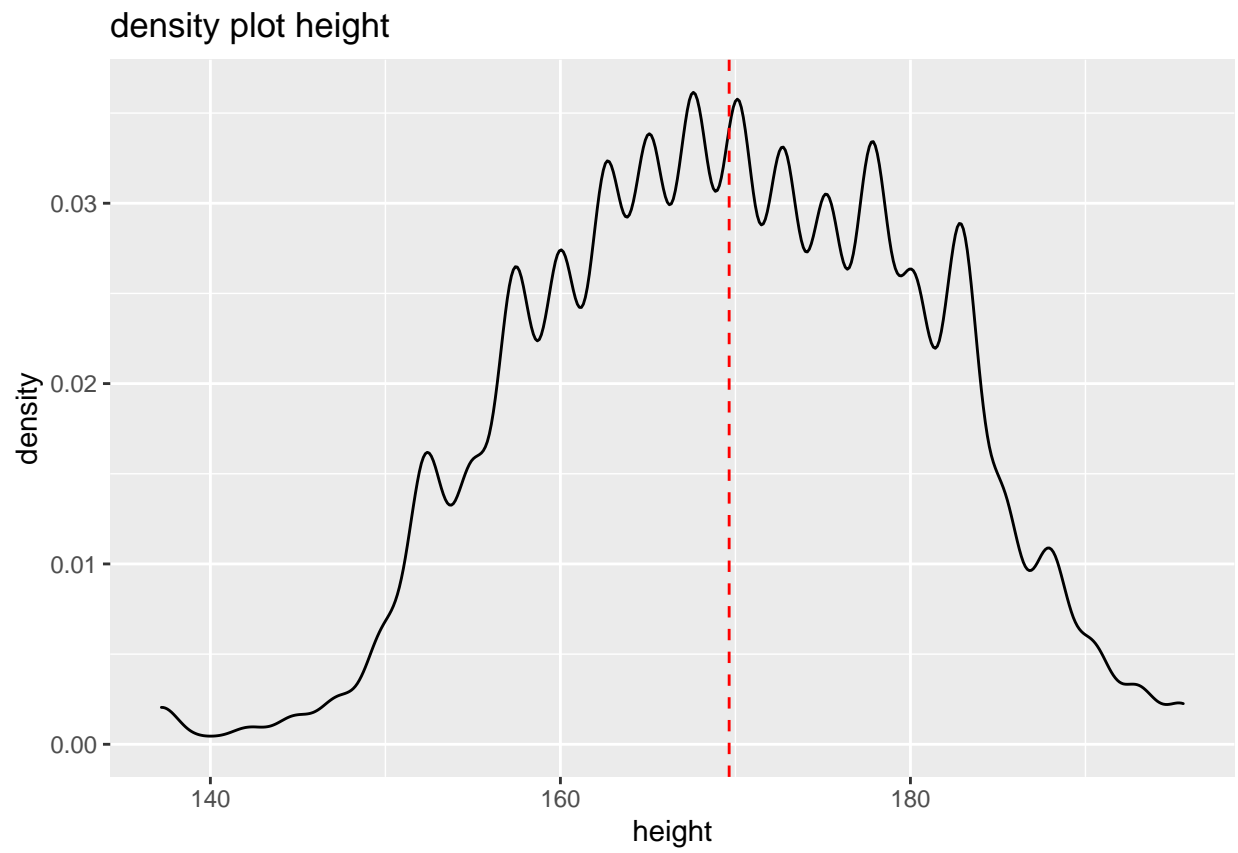
\* Lebih banyak bersifat continuous column \* ada missing column seluruhnya \* pada masing masing data, ada data yang missing

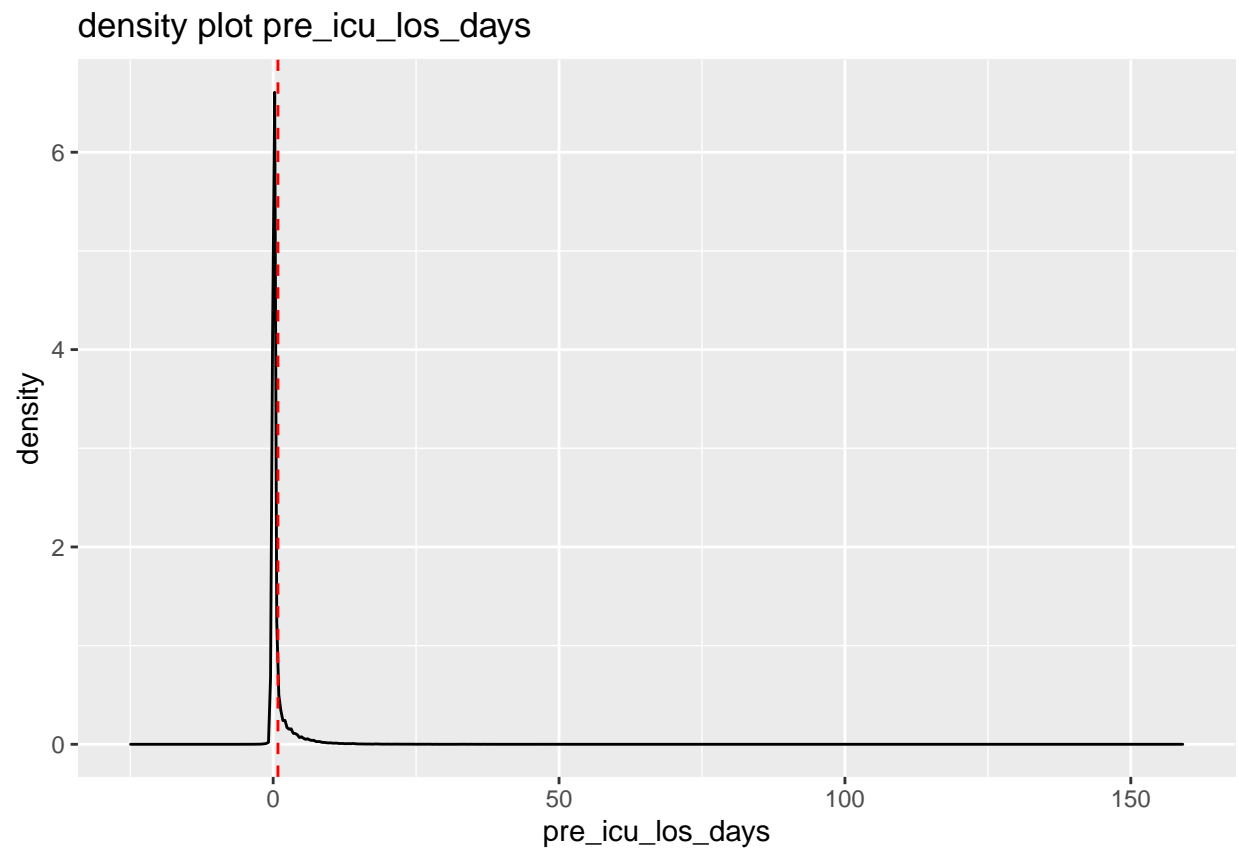
### Density Plot

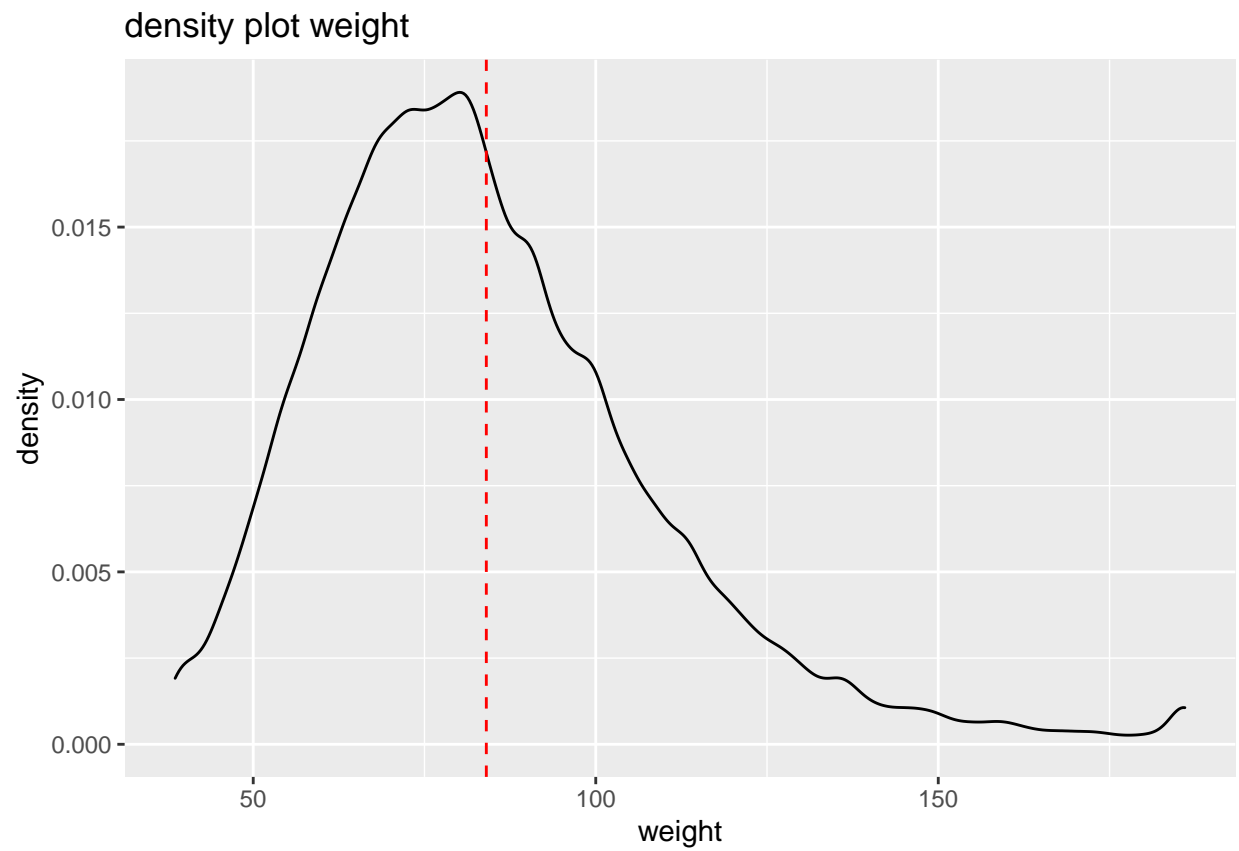
```
numeric_attr <- names(which(lapply(data,class)=="numeric"))
for(i in numeric_attr){
  print(ggplot(data,aes_string(x=i)) +
    geom_density() +
    geom_vline(xintercept = mean(data[[i]],na.rm = TRUE),color="red",
      linetype="dashed" ) +
    ggtitle(paste("density plot ",i,sep="")))
}
```

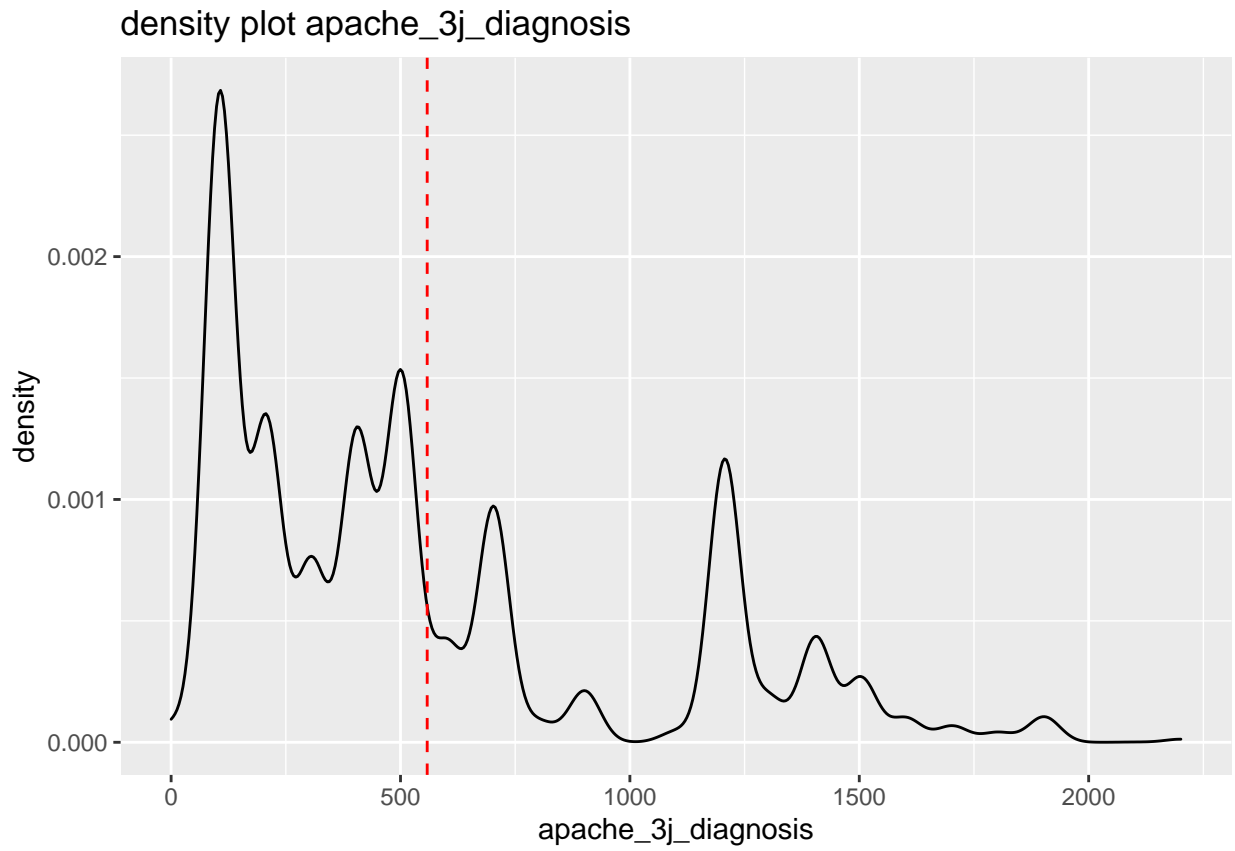


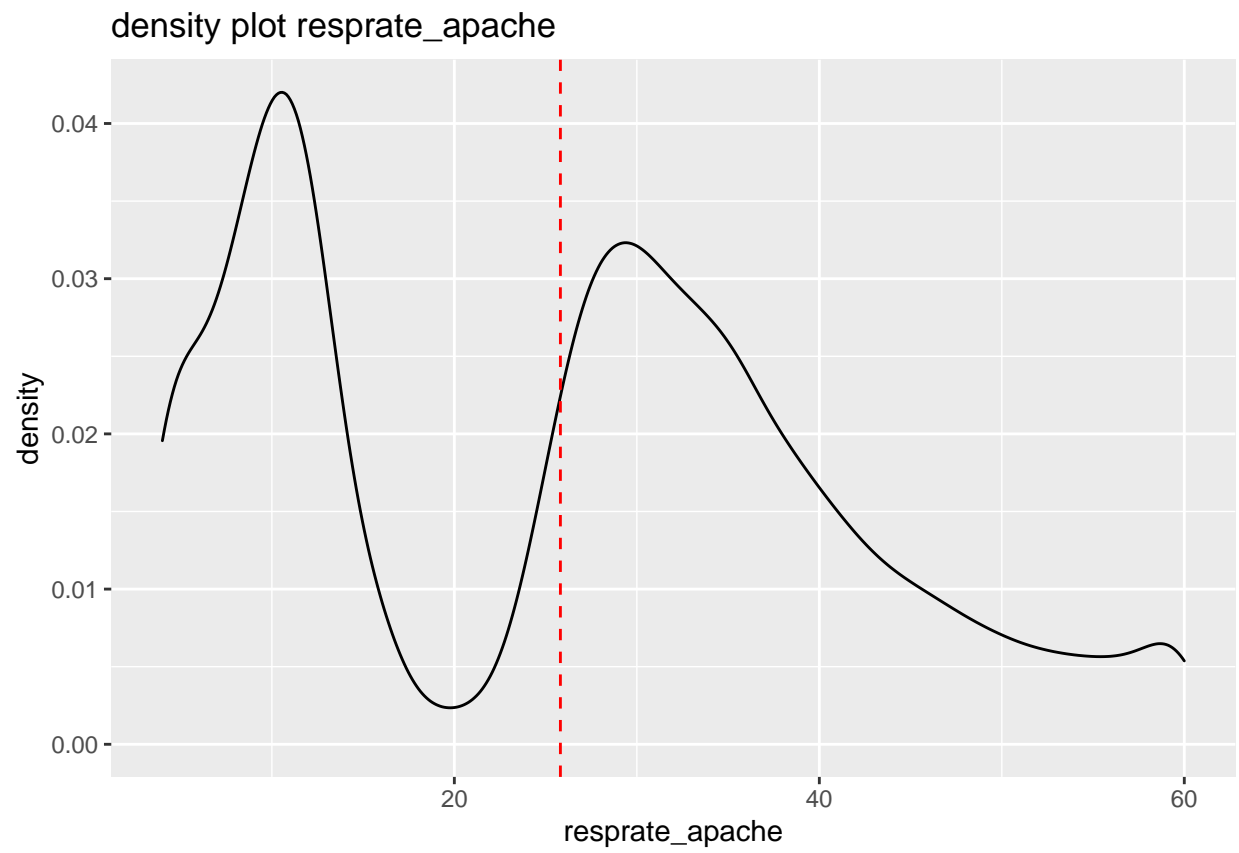


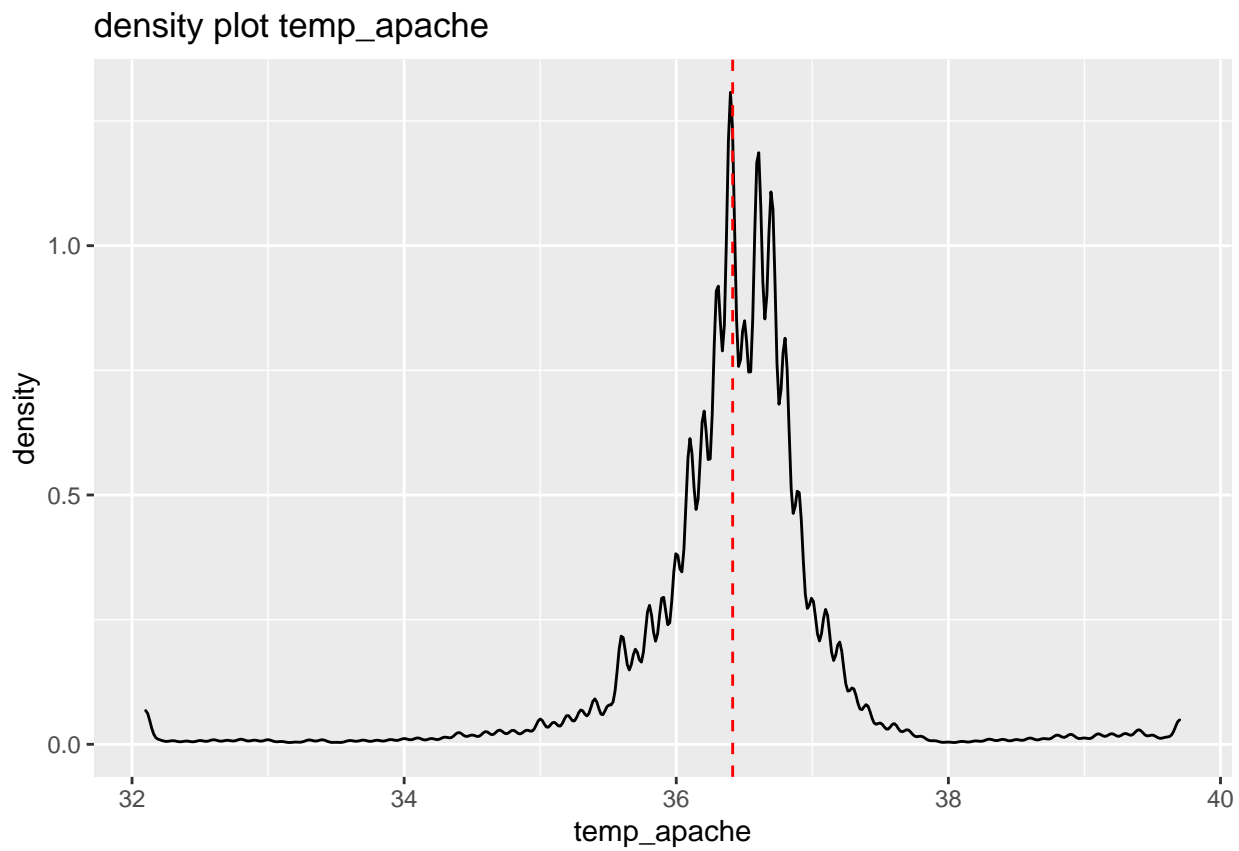


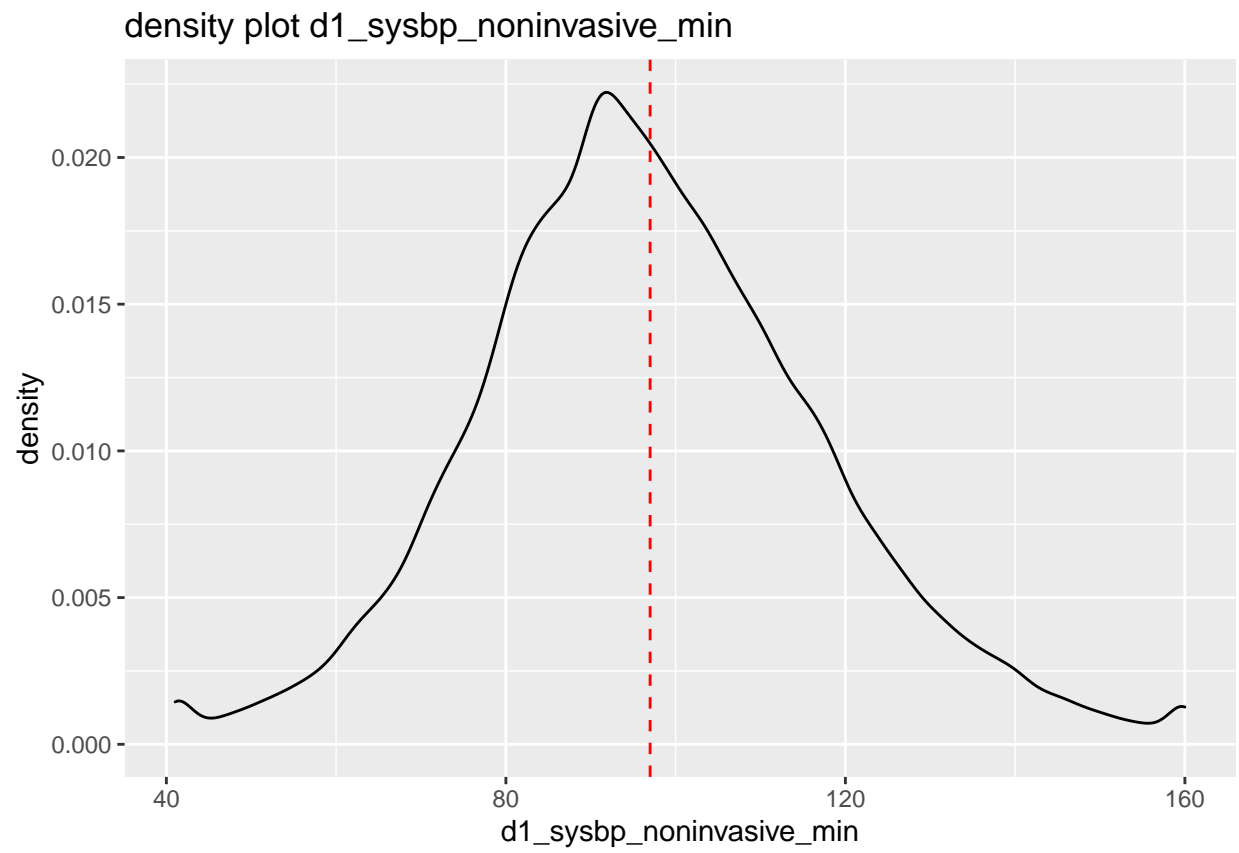


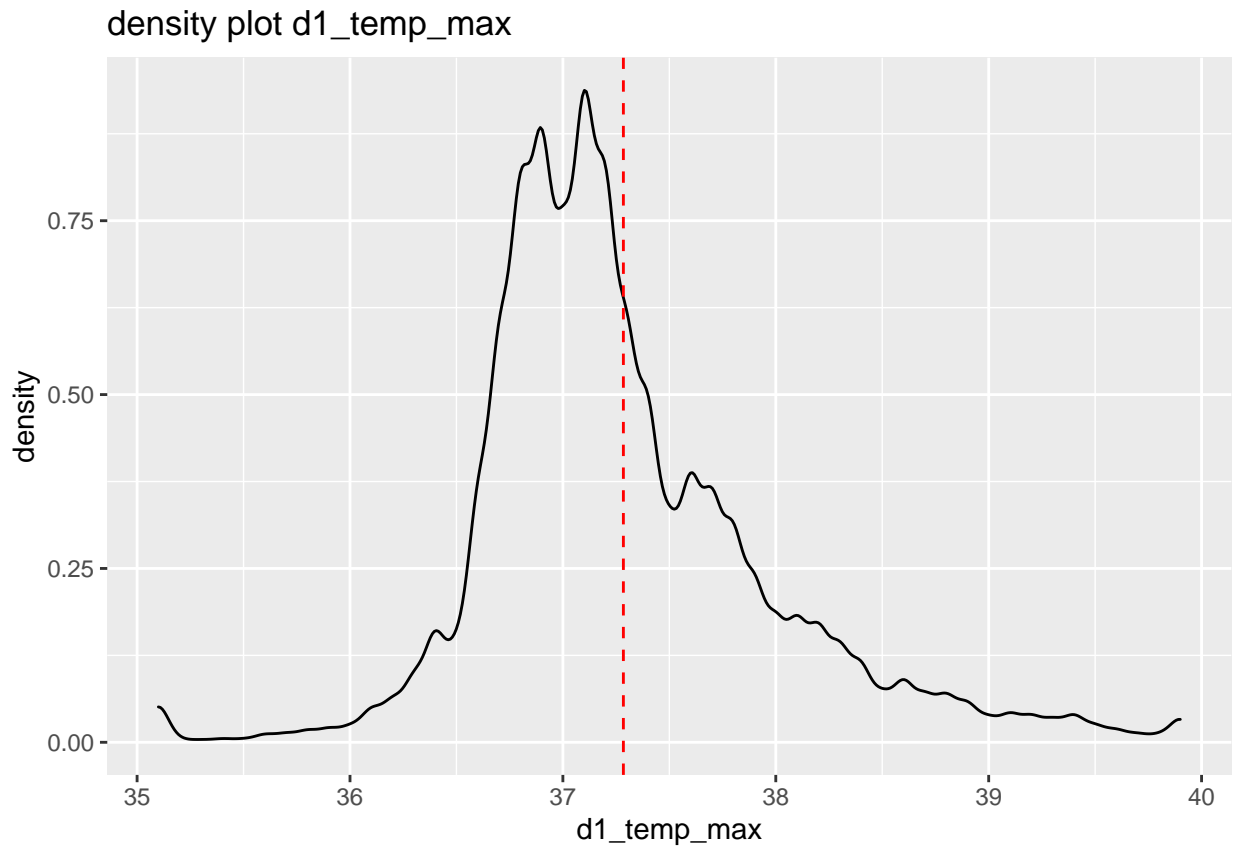




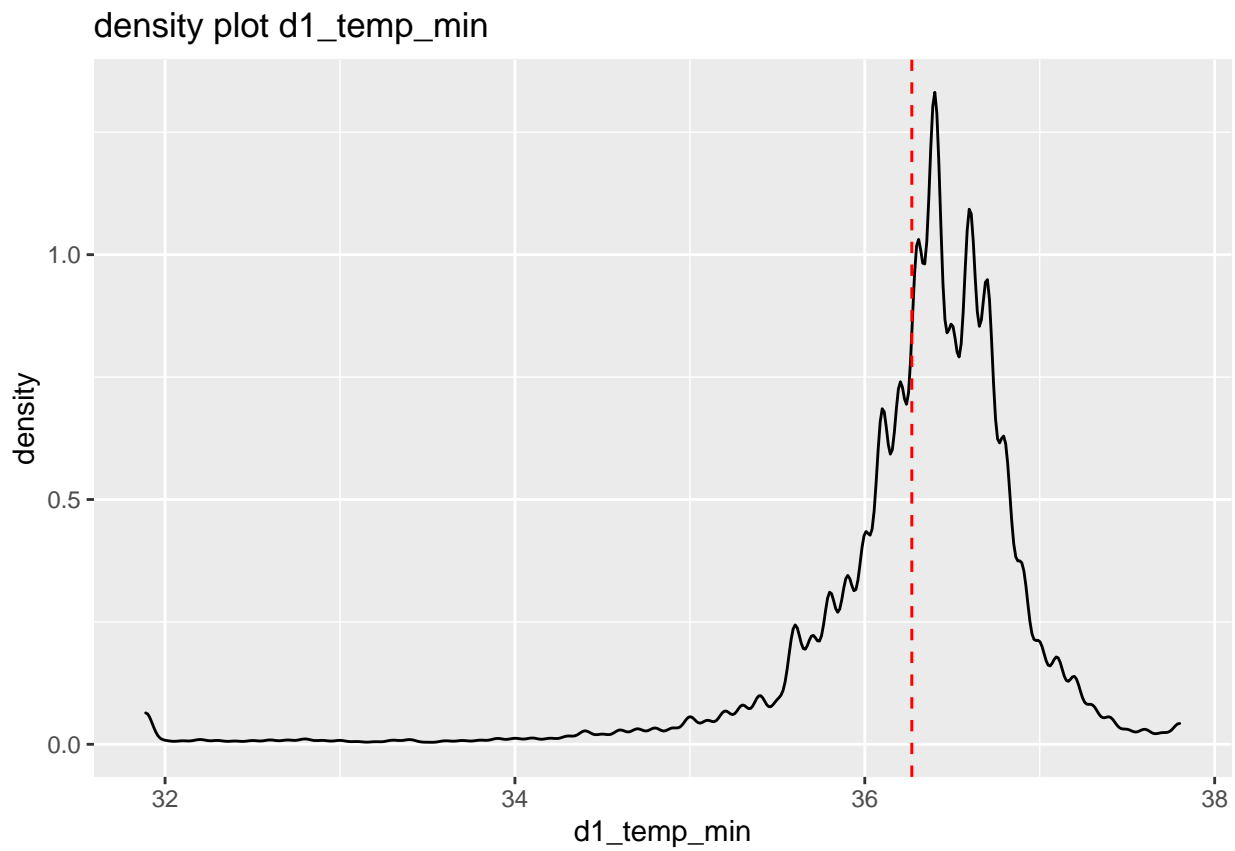


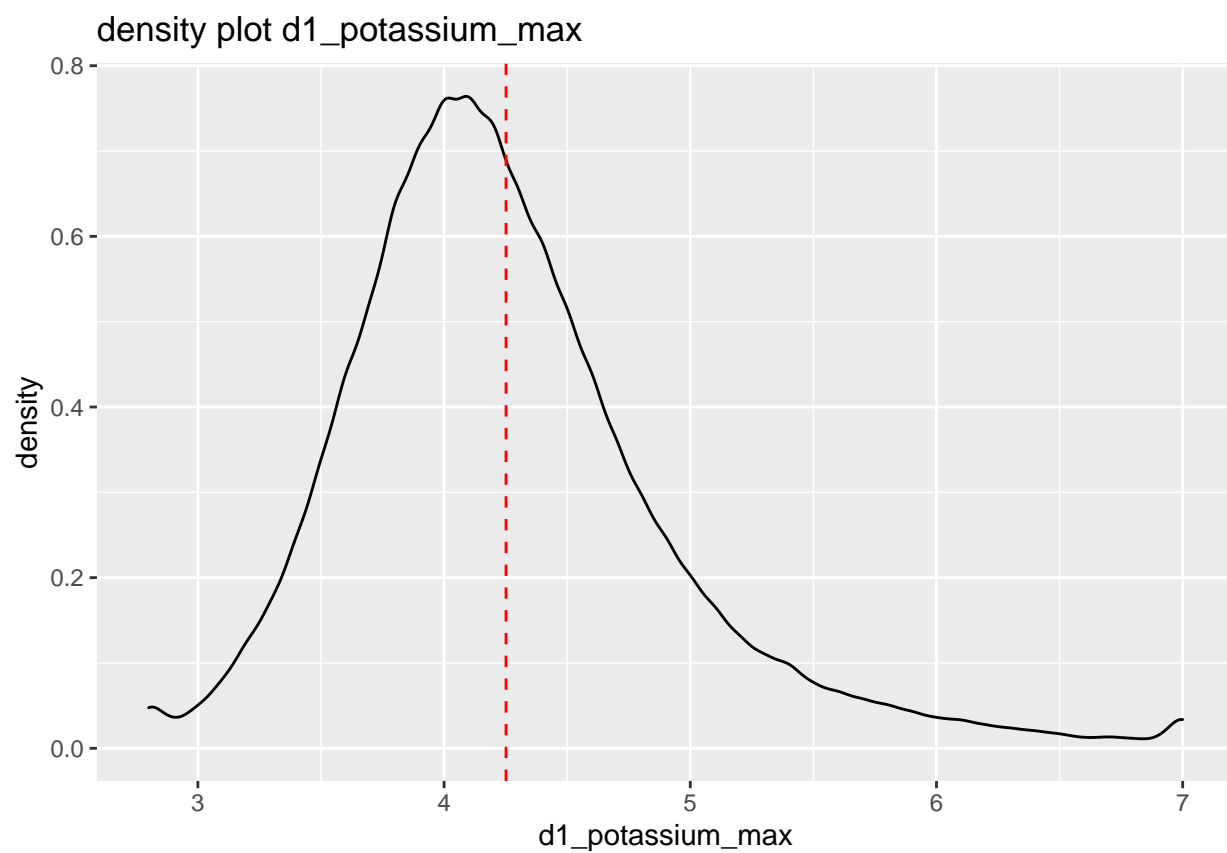


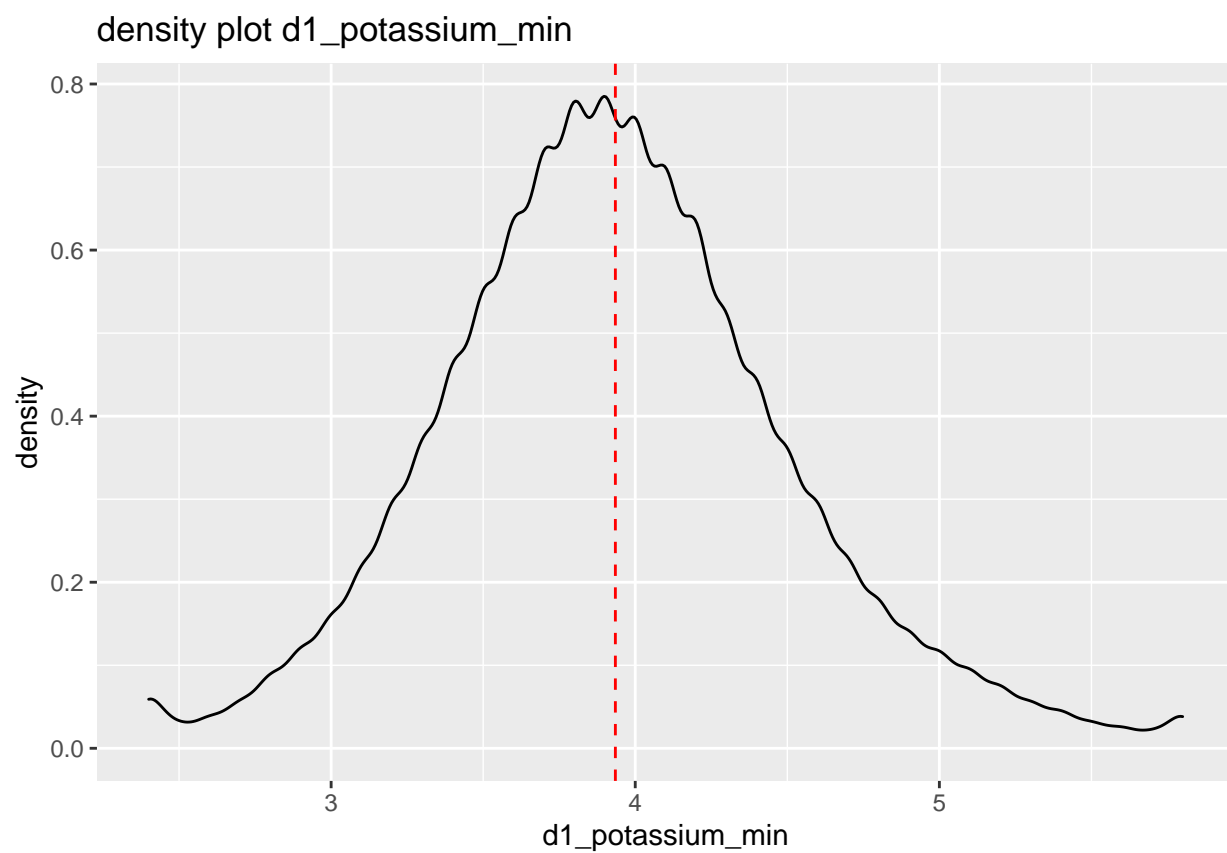


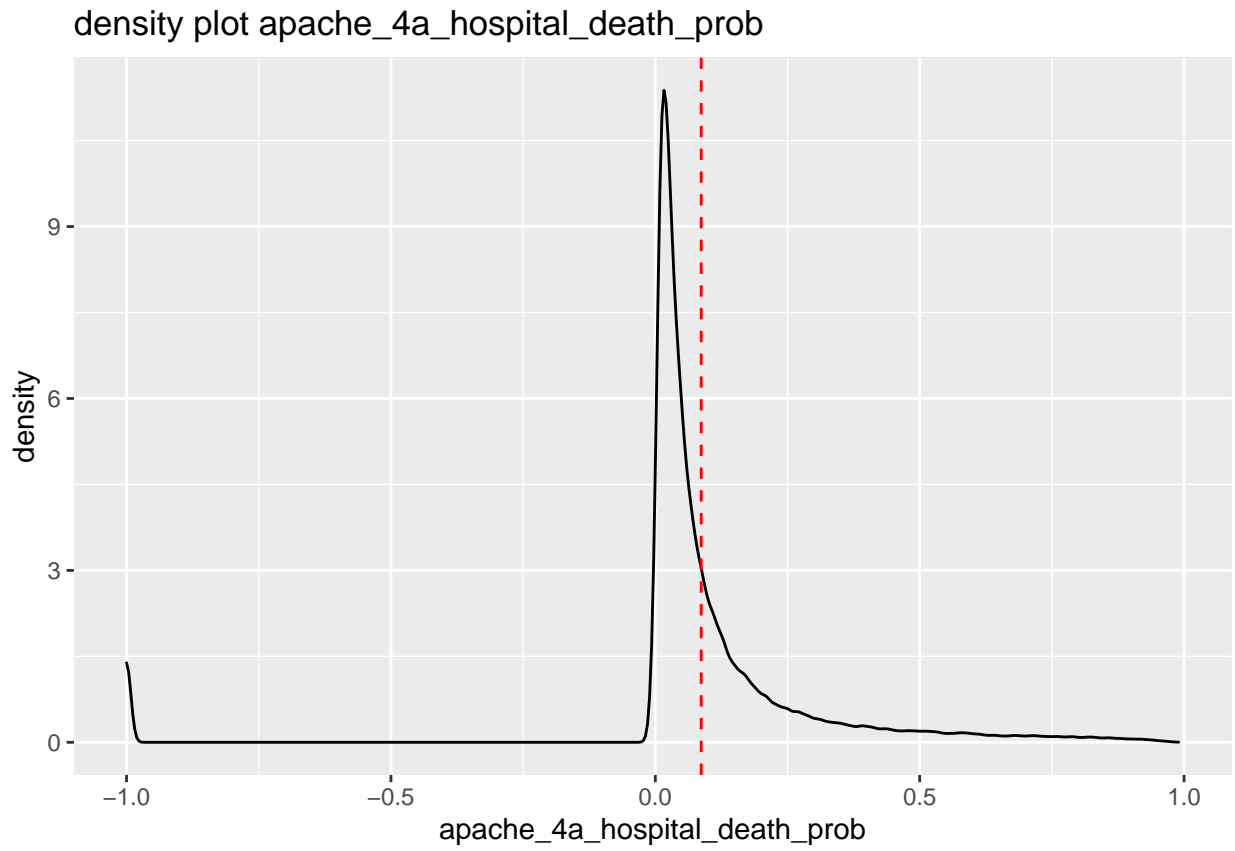


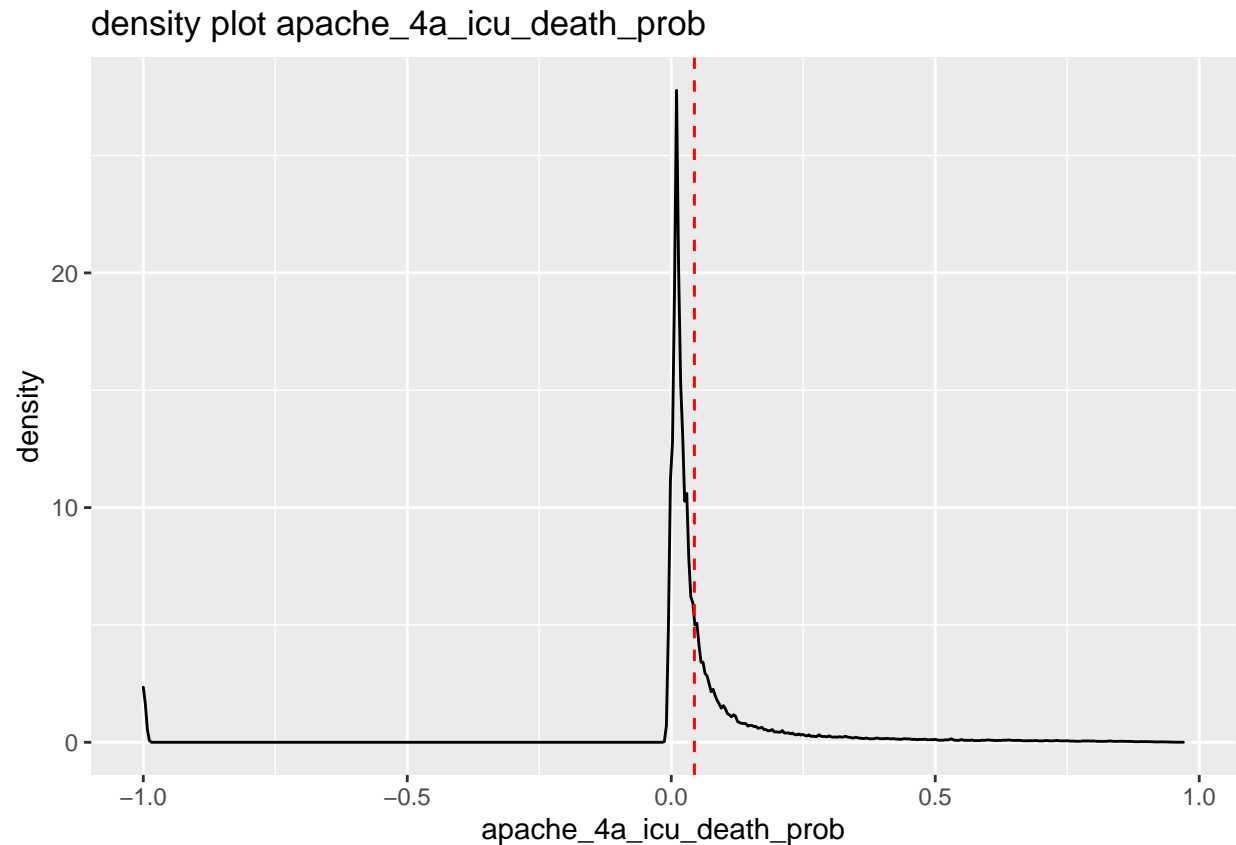








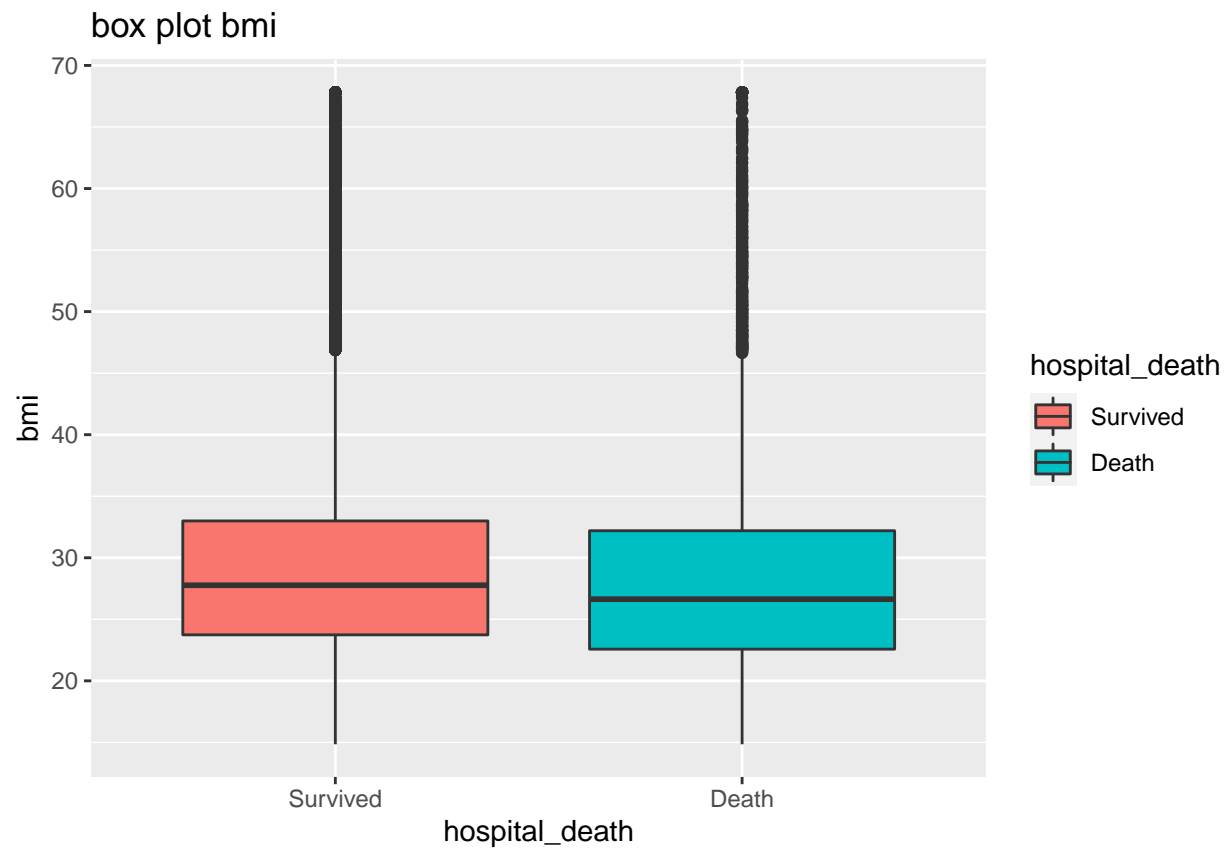


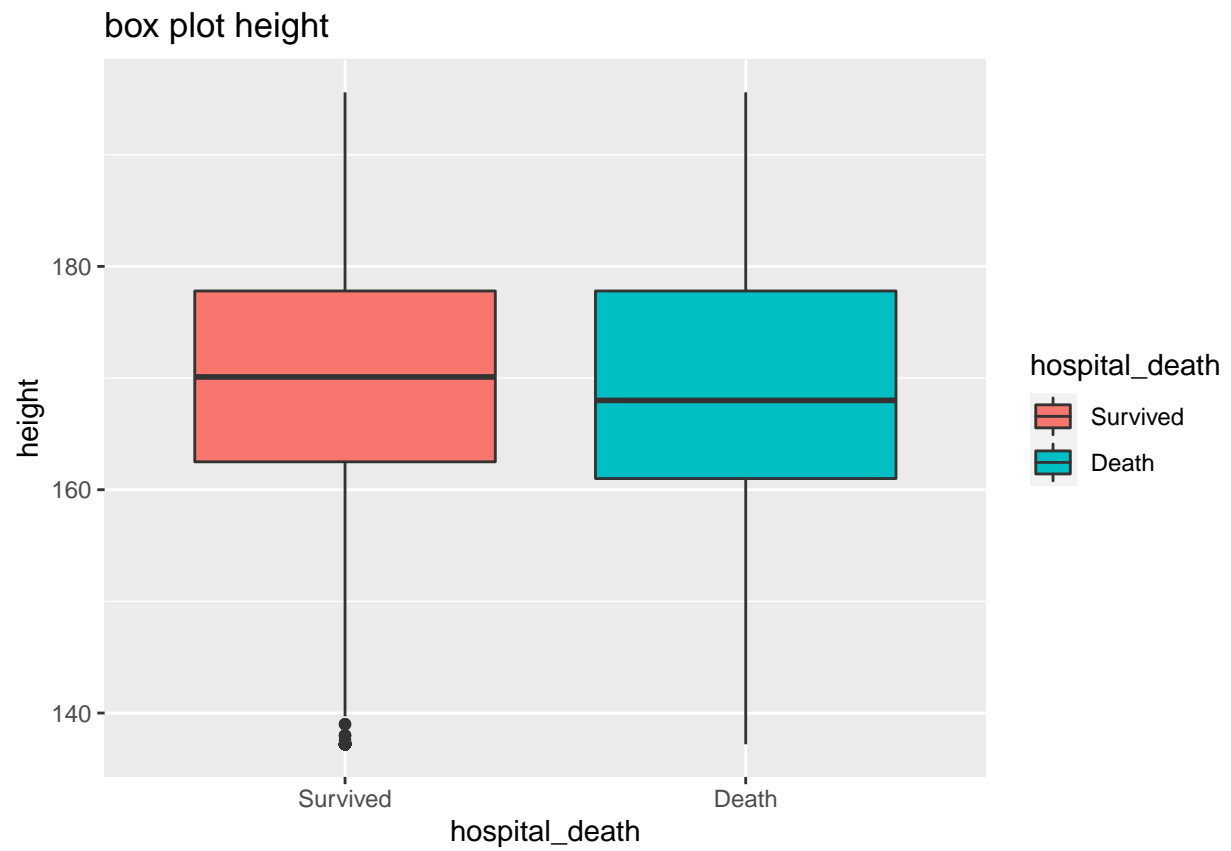


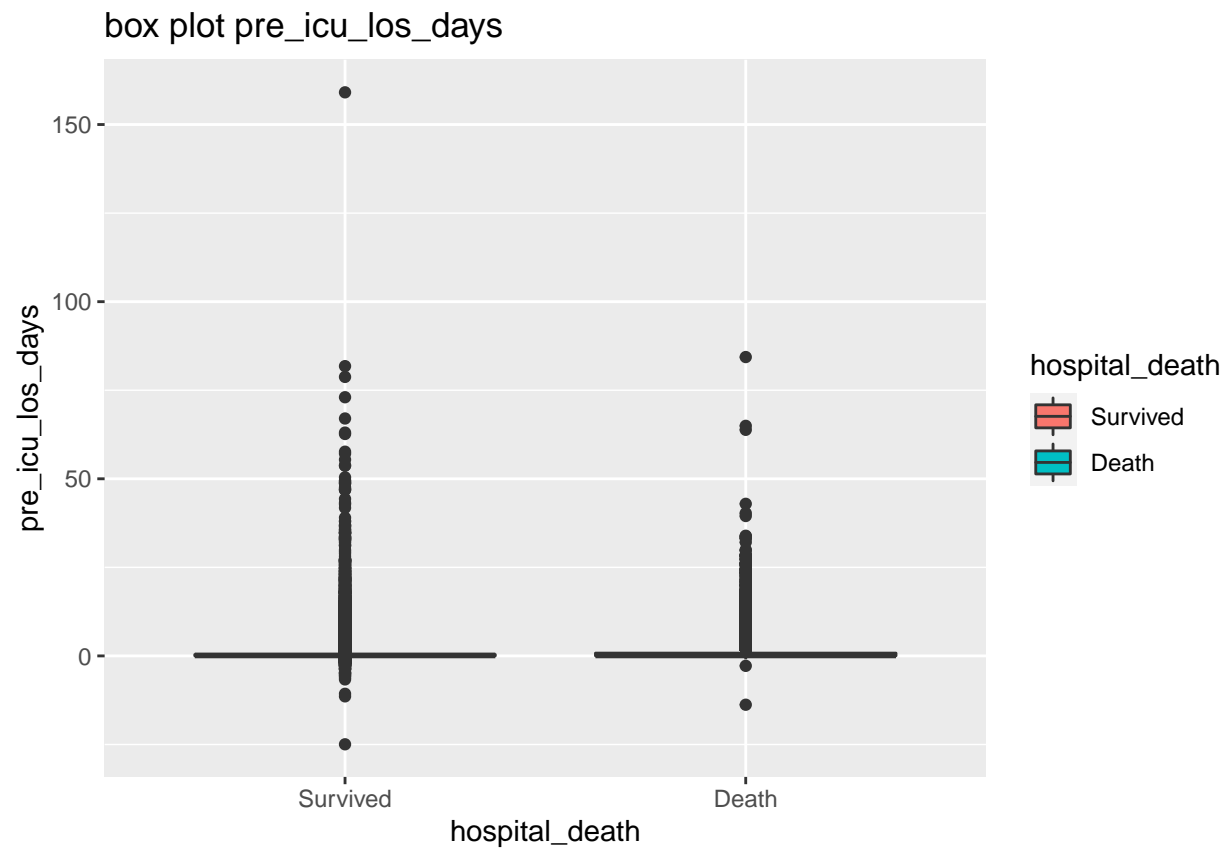
\* Jika Diperhatikan, Banyak Sekali Data yang puncak densitynya tidak cocok dengan mean \* Cenderung menjulur ke arah kanan maupun kiri, \* Distribusi tidak Normal \* Kesimpulannya lebih baik digunakan **imputasi Median** dibandingkan **mean**

### Boxplot

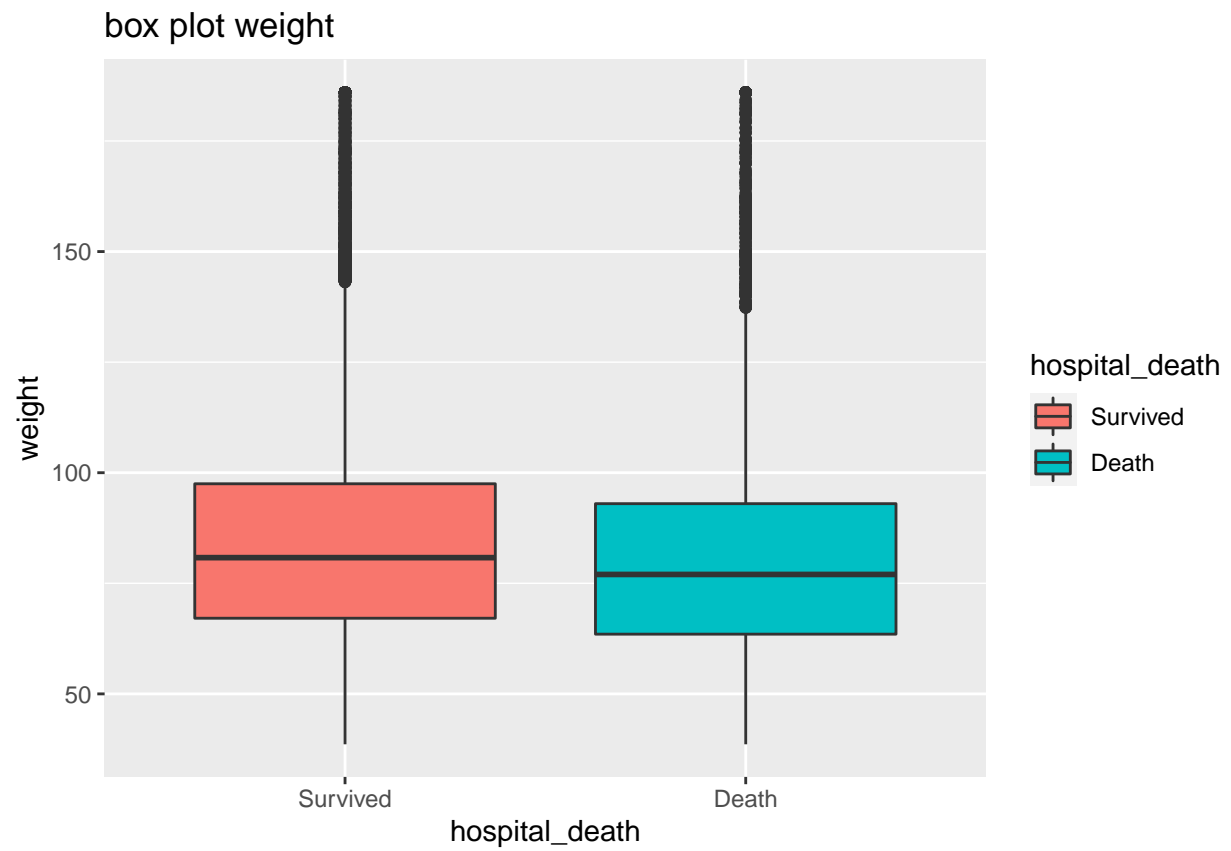
```
for(i in numeric_attr){
  print(ggplot(data,aes_string(x="hospital_death",y=i,fill="hospital_death")) +
    geom_boxplot() +
    ggtitle(paste("box plot ",i,sep="")))
}
```

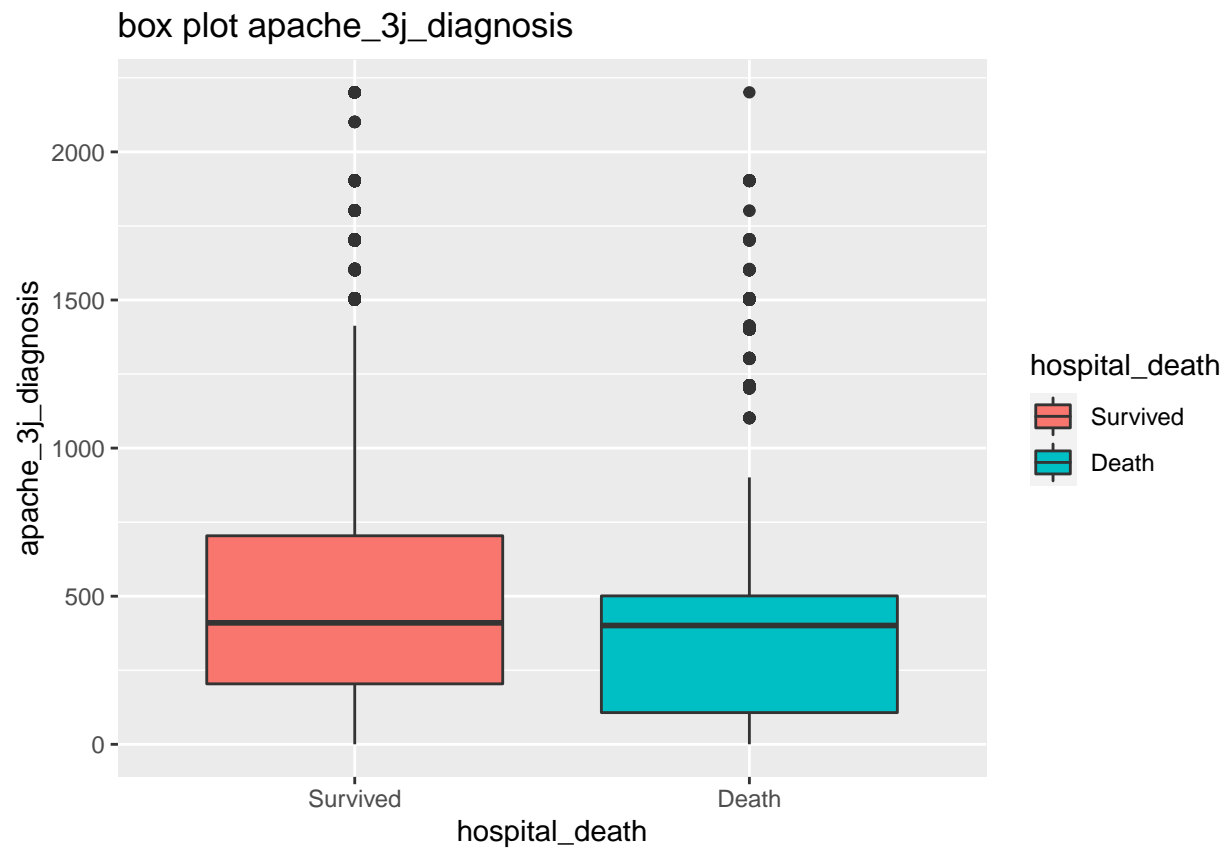


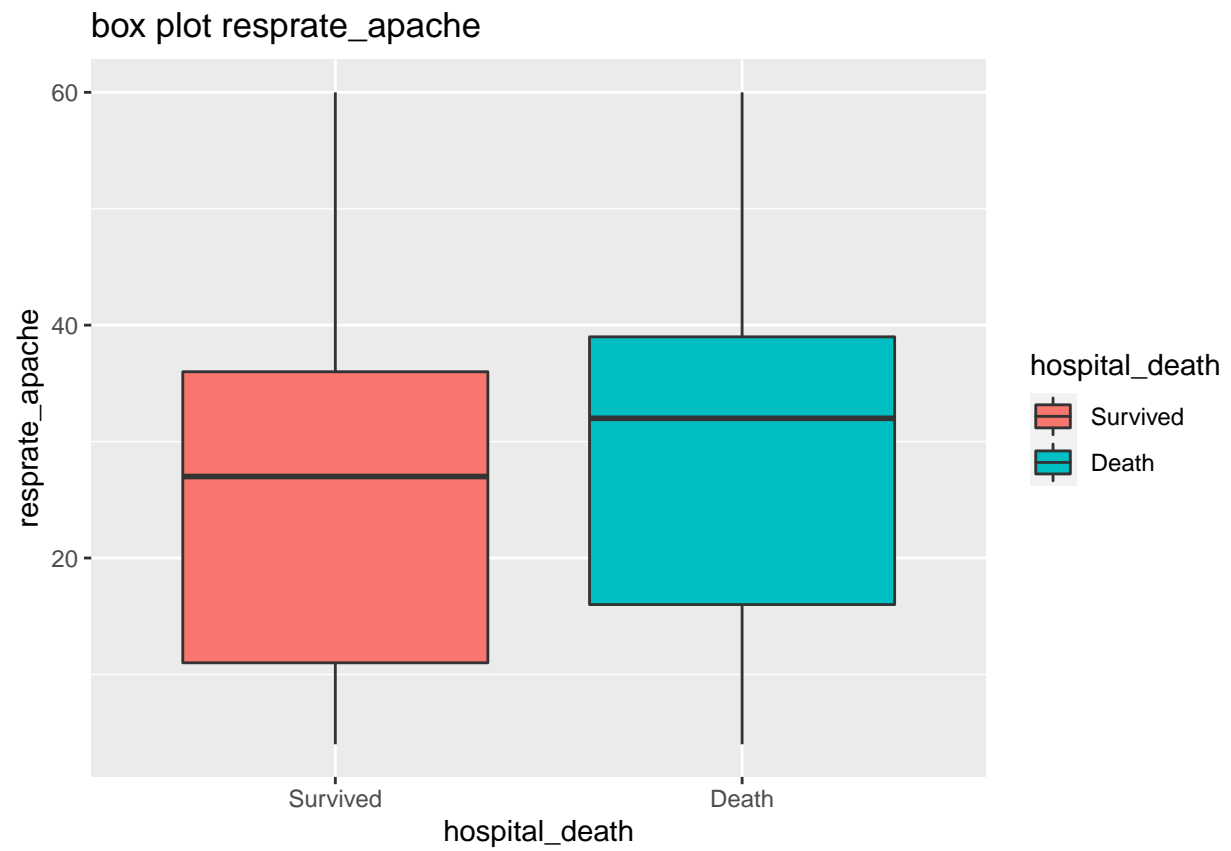


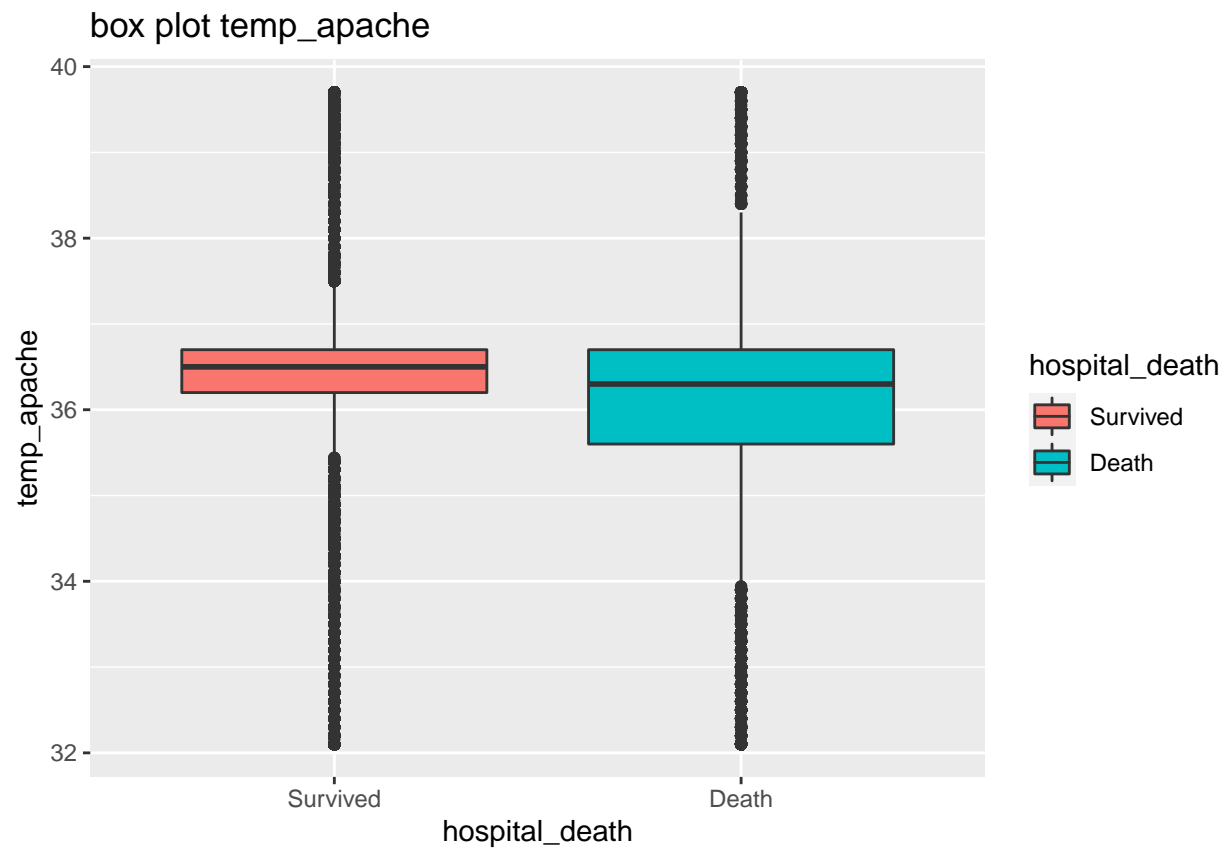


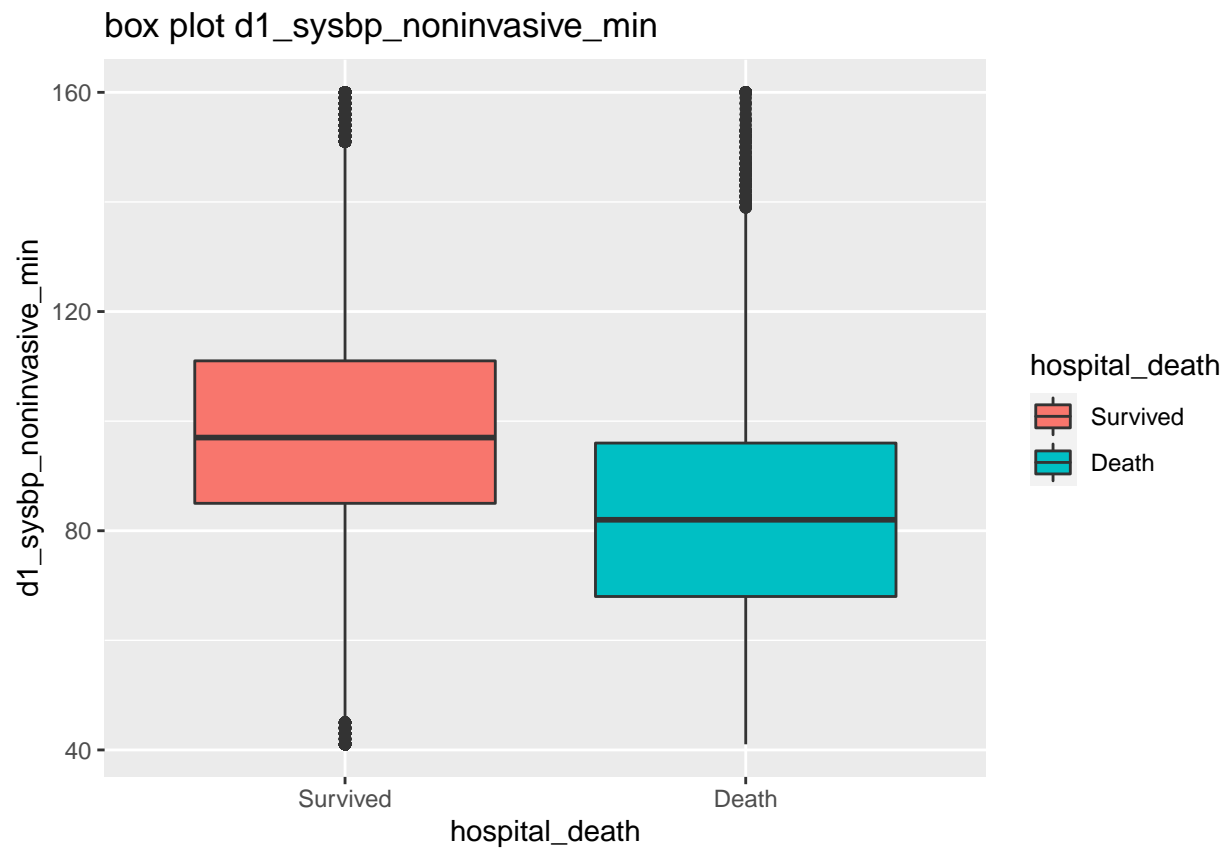


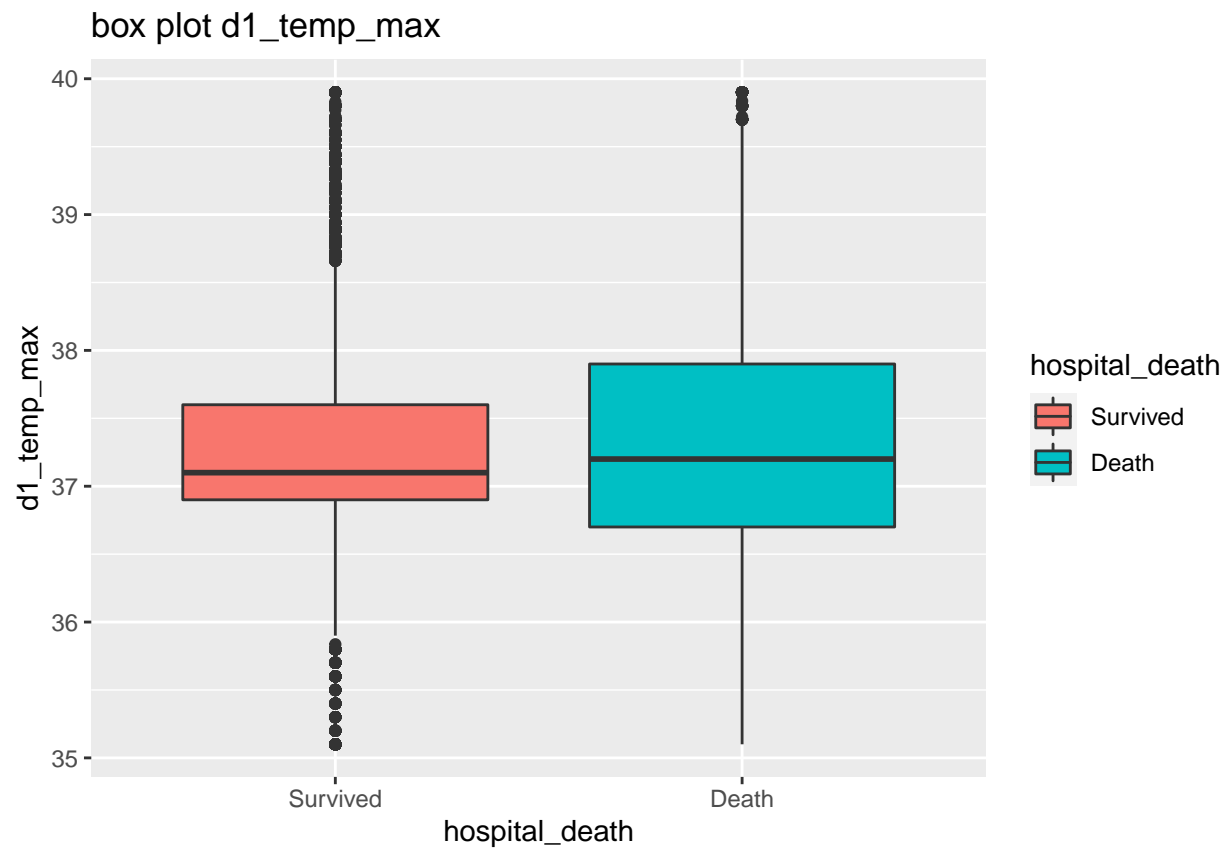


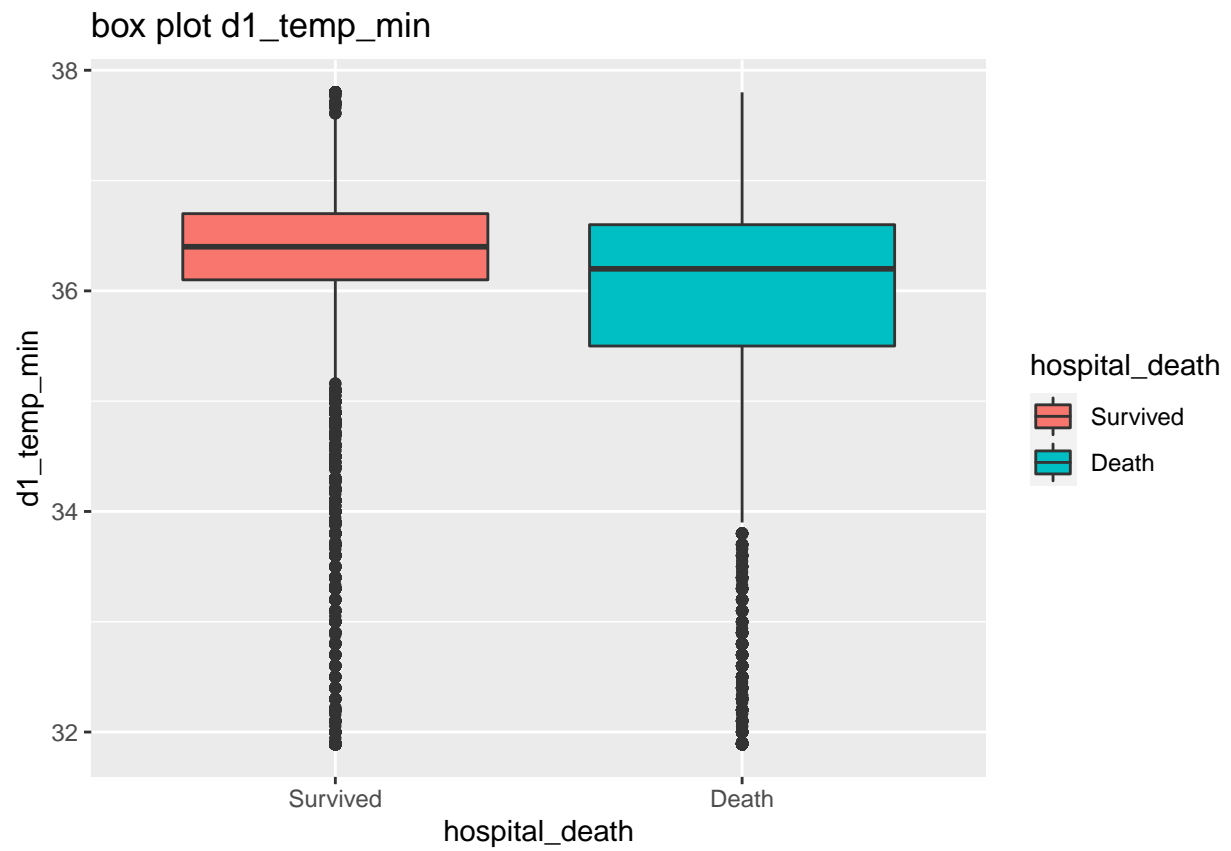


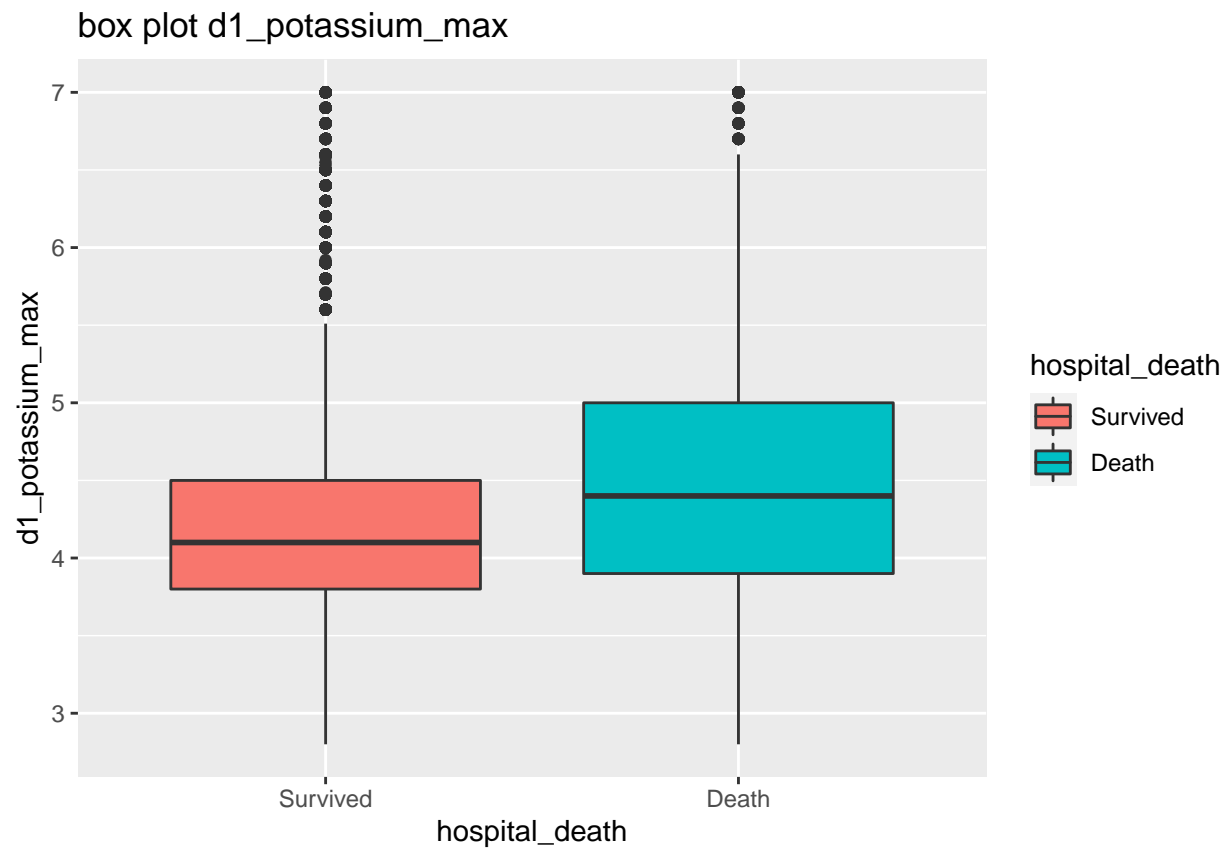




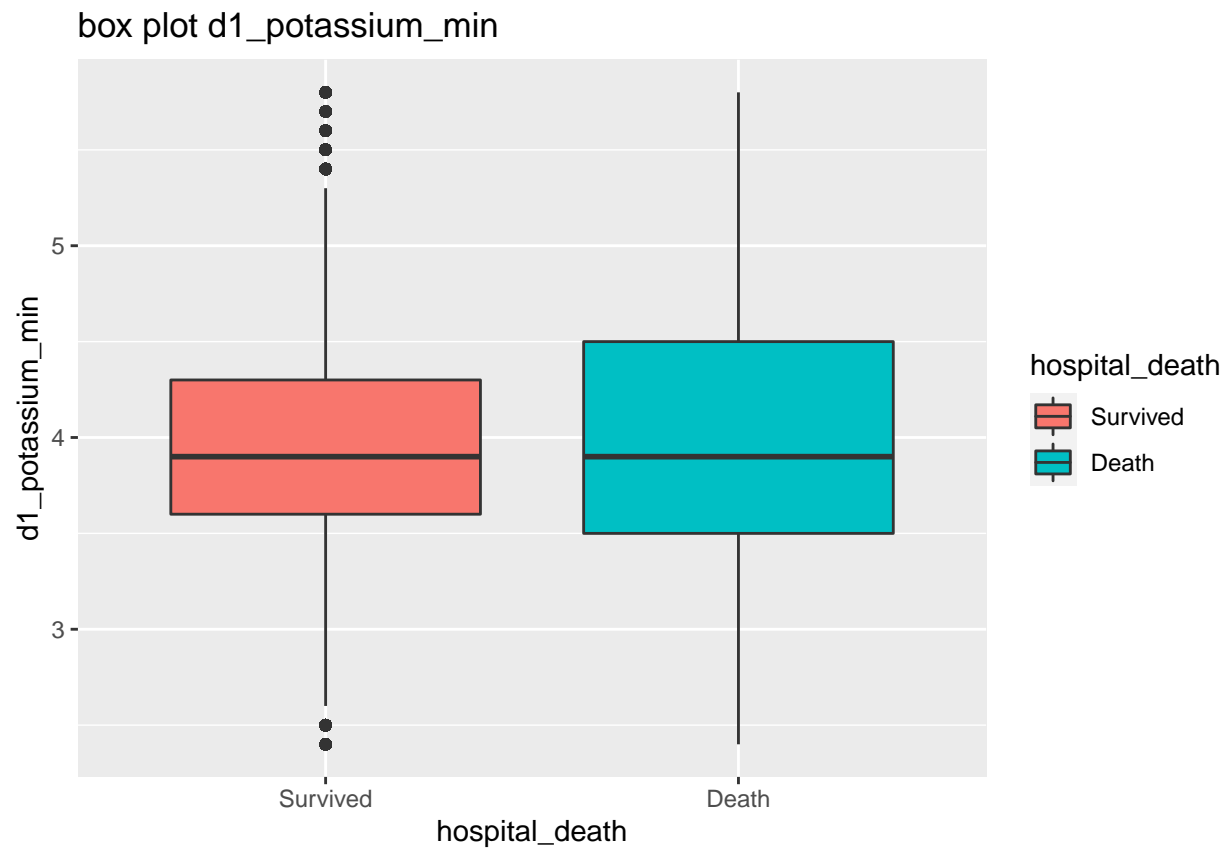


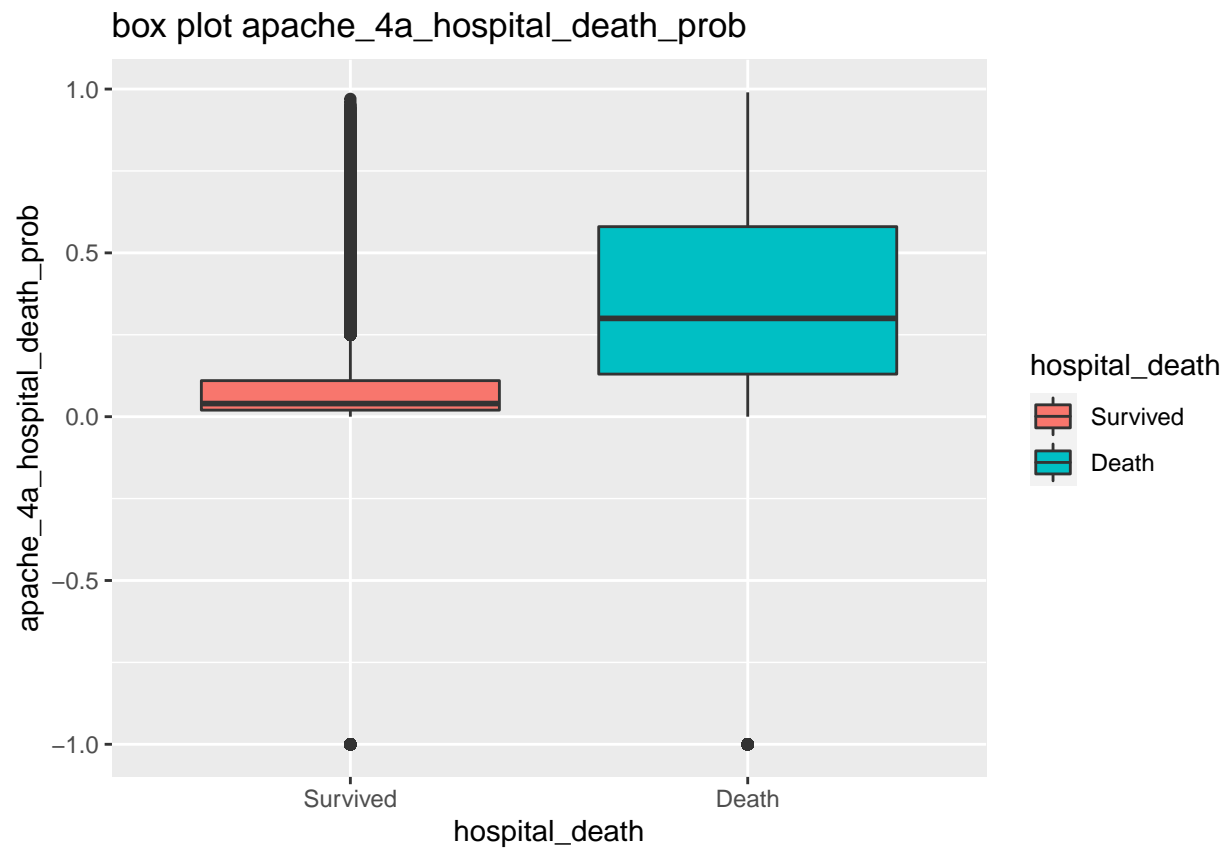


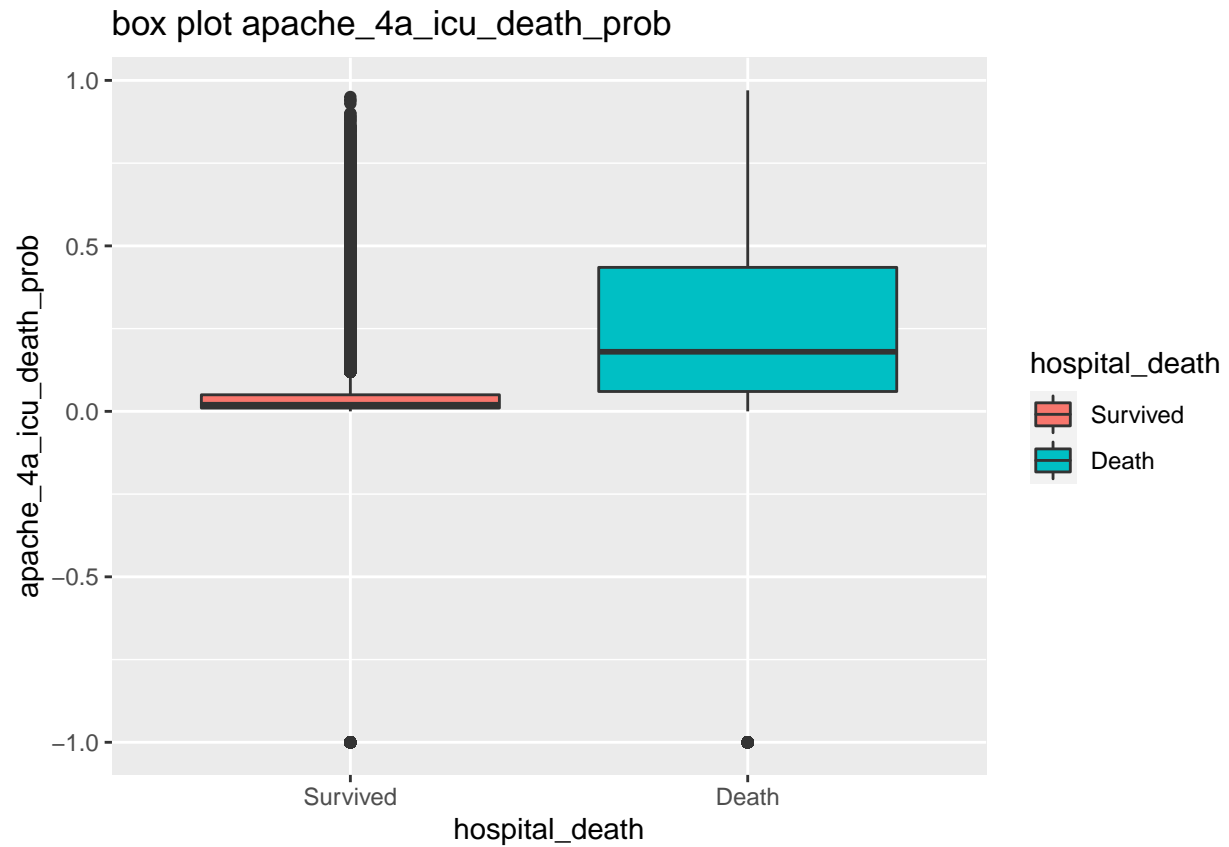








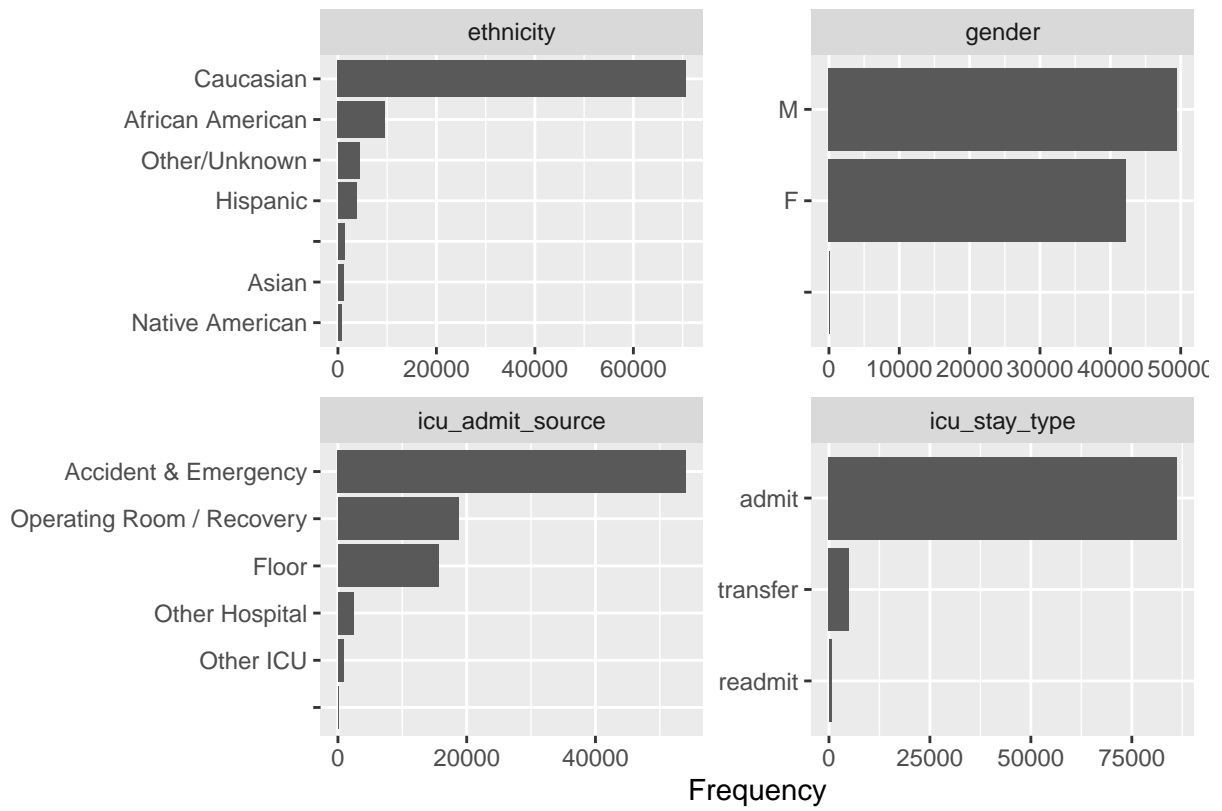


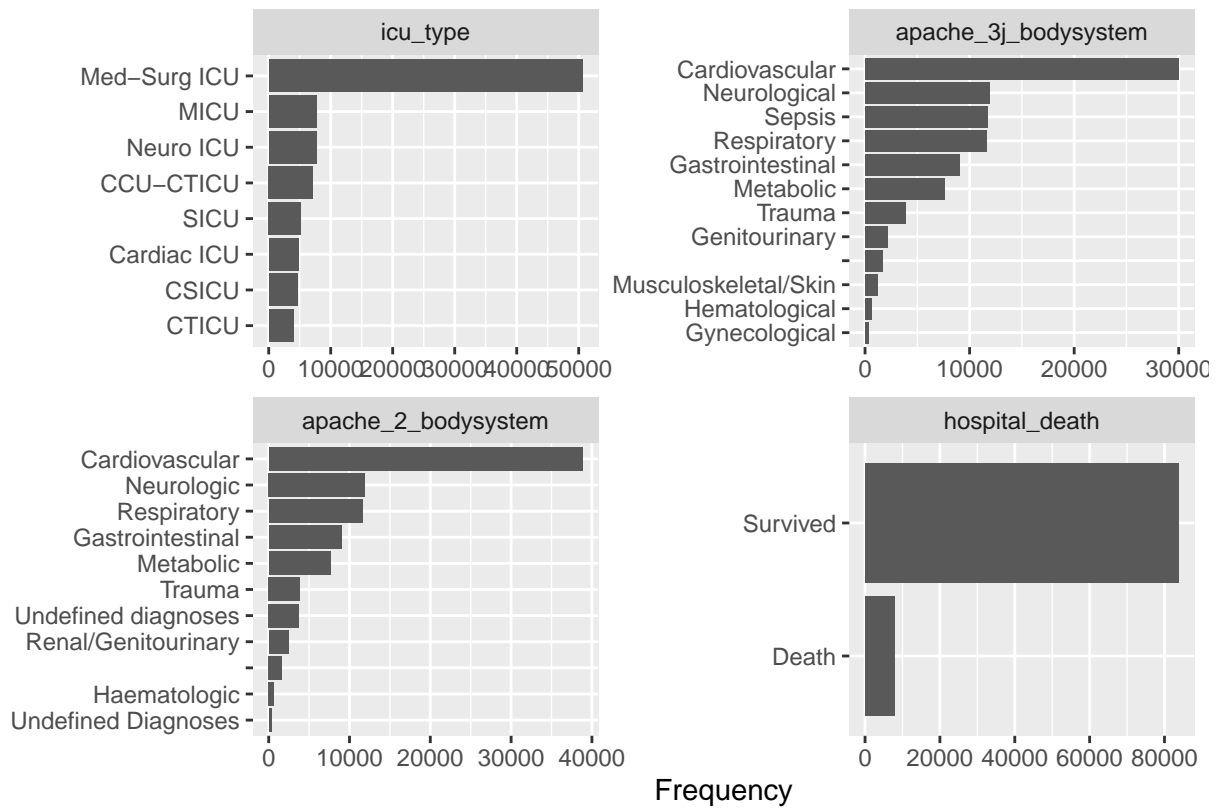


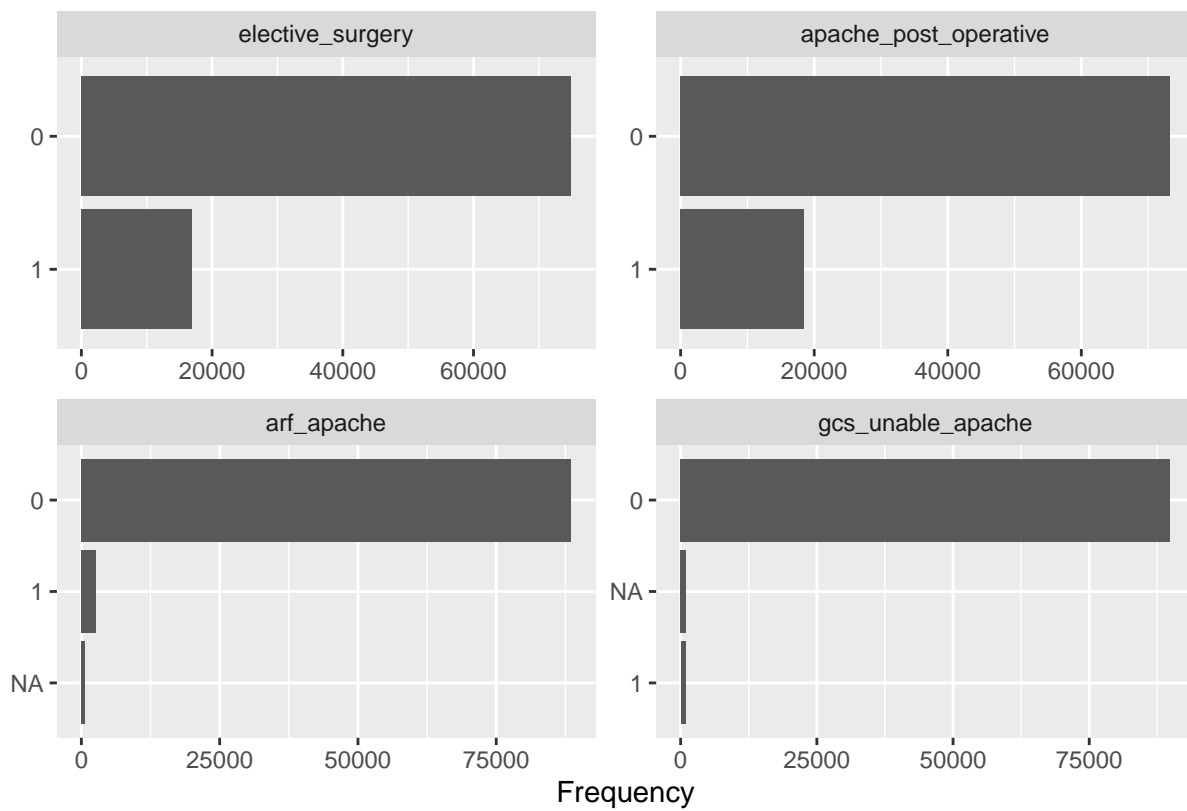
\* Jika Diperhatikan, Keseluruhan memiliki outlier . maka sudah dipastikan imputasi median Yang terbaik

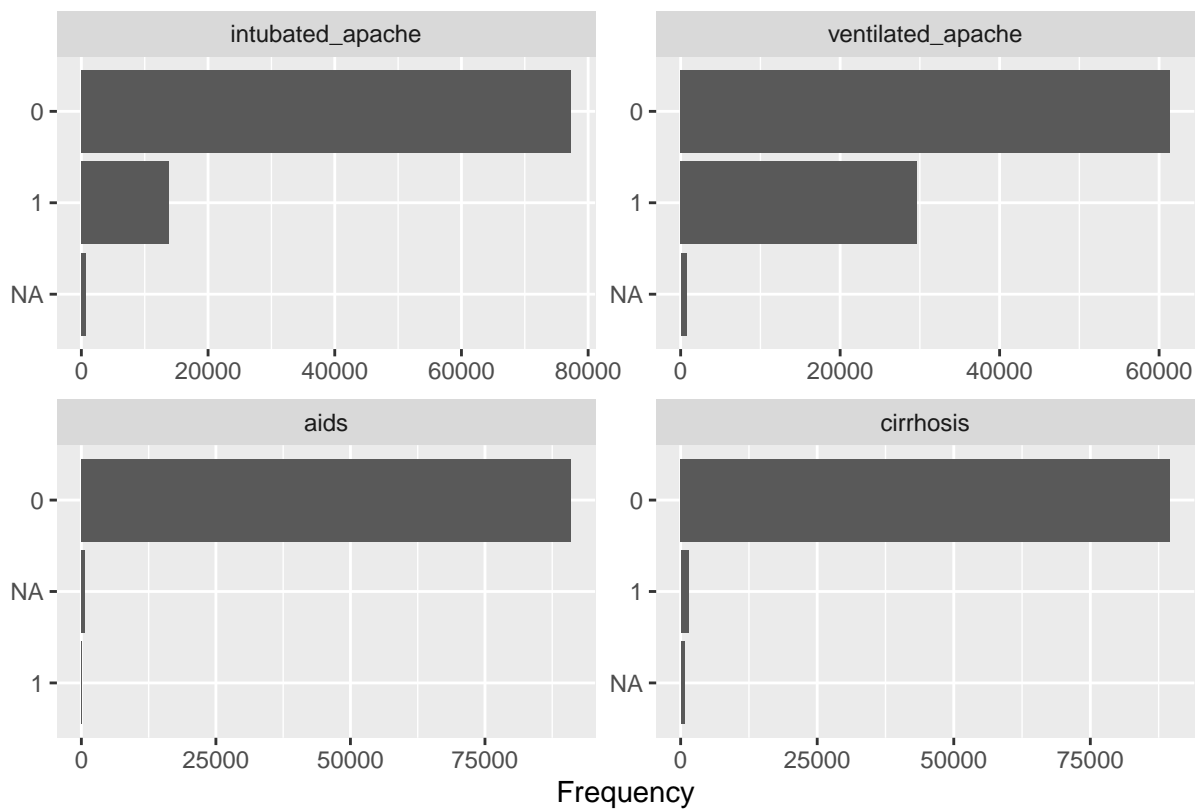
### Categoric Explore

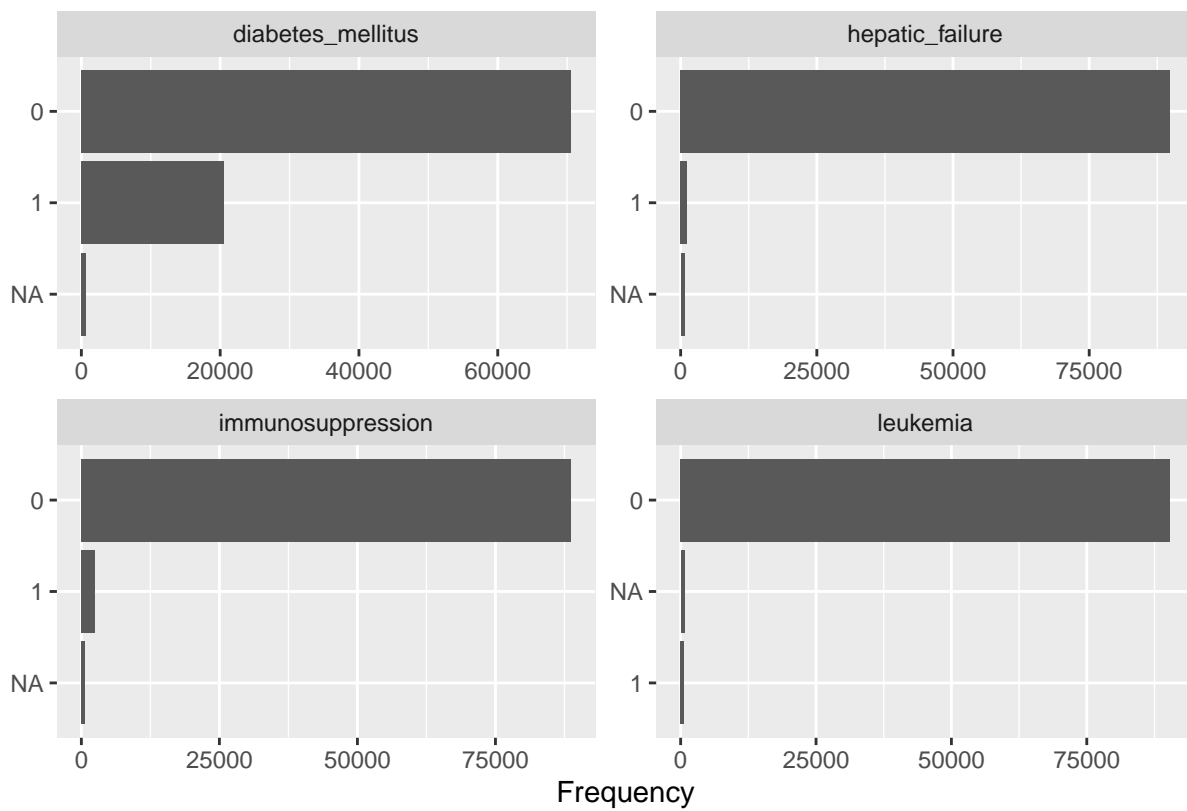
```
DataExplorer::plot_bar(data,nrow=2,ncol=2)
```



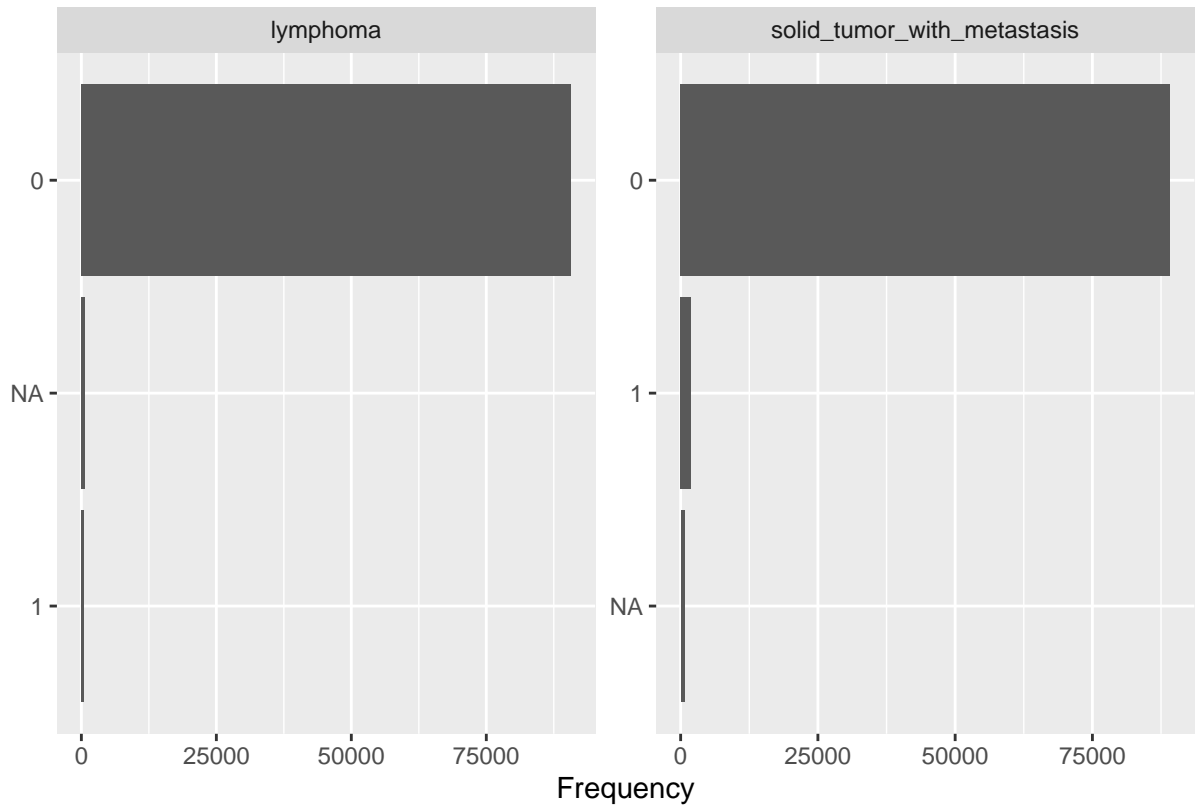










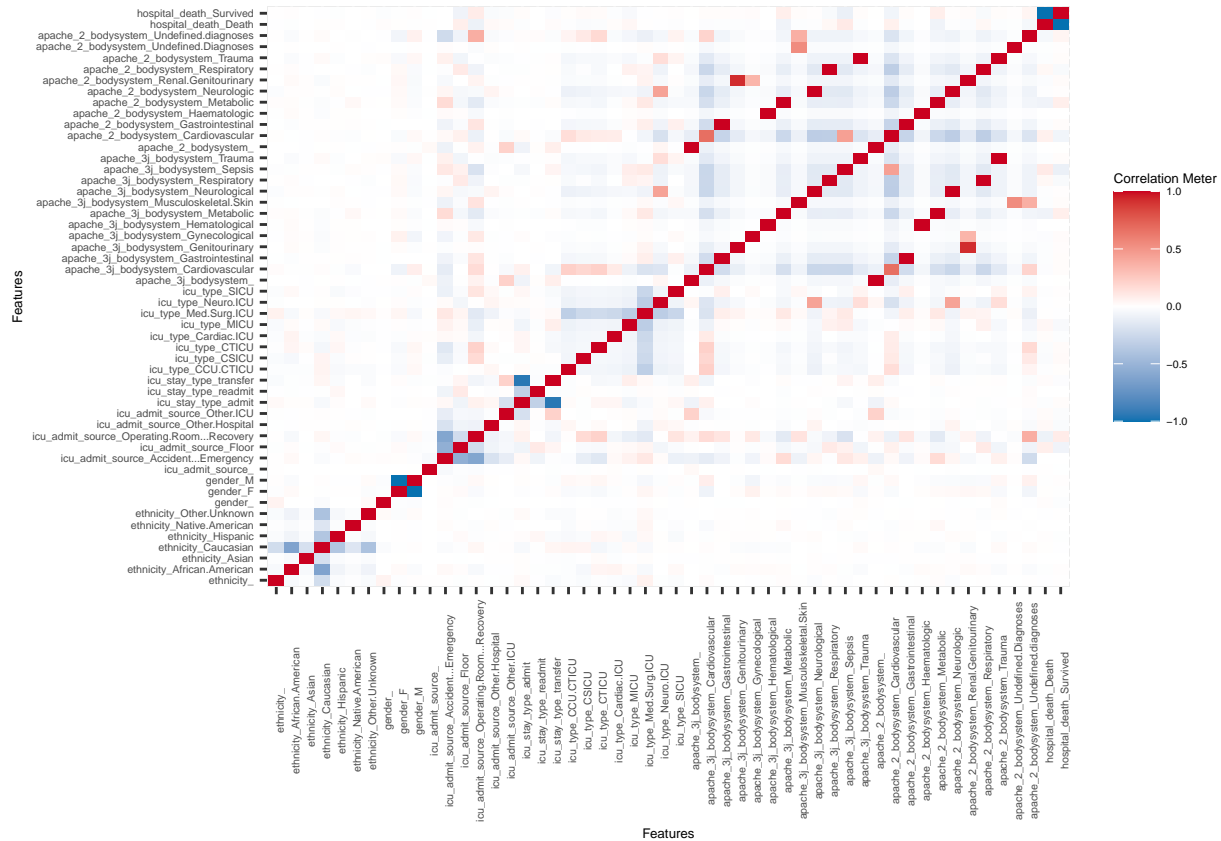


Page 6

\* Masih Ada NA

### Plot correlation for discrete

```
plot_correlation(data,type="discrete",
                 theme_config = list(text=element_text(size=5),
                                     axis.text.x=element_text(angle=90)))
```



- Hanya Sedikit yang berkorelasi ya . meski begitu, ada kemungkinan multikolinearitas disana ya

## Handling Missing Value

### Check Table frequency NA

```
# Check column if have NA's
kolom_NA <- data.frame(as.list(colSums(is.na(data))))
kolom_NA <- as.data.frame(t(kolom_NA))
kolom_NA <- data.frame(names = row.names(kolom_NA), kolom_NA)
rownames(kolom_NA) <- NULL
colnames(kolom_NA) <- c("Nama_Kolom", "freq")

# pipelining https://raw.githubusercontent.com/rstudio/cheatsheets/main/data-transformation.pdf
kolom_NA <- kolom_NA %>%
  mutate(percen_freq = freq / ukuran_data[1] * 100) %>%
  filter(freq > 0)

kolom_NA
```

```
##          Nama_Kolom  freq percen_freq
## 1             age  4228    4.6100335
## 2             bmi  3429    3.7388375
```

## 3	height	1334	1.4545375
## 4	weight	2720	2.9657737
## 5	apache_2_diagnosis	1662	1.8121749
## 6	apache_3j_diagnosis	1101	1.2004841
## 7	arf_apache	715	0.7796059
## 8	gcs_eyes_apache	1901	2.0727705
## 9	gcs_motor_apache	1901	2.0727705
## 10	gcs_unable_apache	1037	1.1307012
## 11	gcs_verbal_apache	1901	2.0727705
## 12	heart_rate_apache	878	0.9573343
## 13	intubated_apache	715	0.7796059
## 14	map_apache	994	1.0838158
## 15	resprate_apache	1234	1.3455017
## 16	temp_apache	4108	4.4791905
## 17	ventilated_apache	715	0.7796059
## 18	d1_diasbp_max	165	0.1799091
## 19	d1_diasbp_min	165	0.1799091
## 20	d1_diasbp_noninvasive_max	1040	1.1339723
## 21	d1_diasbp_noninvasive_min	1040	1.1339723
## 22	d1_heartrate_max	145	0.1581019
## 23	d1_heartrate_min	145	0.1581019
## 24	d1_mbp_max	220	0.2398788
## 25	d1_mbp_min	220	0.2398788
## 26	d1_mbp_noninvasive_max	1479	1.6126394
## 27	d1_mbp_noninvasive_min	1479	1.6126394
## 28	d1_resprate_max	385	0.4197878
## 29	d1_resprate_min	385	0.4197878
## 30	d1_spo2_max	333	0.3630892
## 31	d1_spo2_min	333	0.3630892
## 32	d1_sysbp_max	159	0.1733669
## 33	d1_sysbp_min	159	0.1733669
## 34	d1_sysbp_noninvasive_max	1027	1.1197976
## 35	d1_sysbp_noninvasive_min	1027	1.1197976
## 36	d1_temp_max	2324	2.5339919
## 37	d1_temp_min	2324	2.5339919
## 38	h1_diasbp_max	3619	3.9460055
## 39	h1_diasbp_min	3619	3.9460055
## 40	h1_diasbp_noninvasive_max	7350	8.0141310
## 41	h1_diasbp_noninvasive_min	7350	8.0141310
## 42	h1_heartrate_max	2790	3.0420987
## 43	h1_heartrate_min	2790	3.0420987
## 44	h1_mbp_max	4639	5.0581706
## 45	h1_mbp_min	4639	5.0581706
## 46	h1_mbp_noninvasive_max	9084	9.9048117
## 47	h1_mbp_noninvasive_min	9084	9.9048117
## 48	h1_resprate_max	4357	4.7506897
## 49	h1_resprate_min	4357	4.7506897
## 50	h1_spo2_max	4185	4.5631481
## 51	h1_spo2_min	4185	4.5631481
## 52	h1_sysbp_max	3611	3.9372826
## 53	h1_sysbp_min	3611	3.9372826
## 54	h1_sysbp_noninvasive_max	7341	8.0043178
## 55	h1_sysbp_noninvasive_min	7341	8.0043178
## 56	d1_glucose_max	5807	6.3317087

```
## 57          d1_glucose_min 5807 6.3317087
## 58          d1_potassium_max 9585 10.4510811
## 59          d1_potassium_min 9585 10.4510811
## 60 apache_4a_hospital_death_prob 7947 8.6650747
## 61      apache_4a_icu_death_prob 7947 8.6650747
## 62              aids 715 0.7796059
## 63              cirrhosis 715 0.7796059
## 64          diabetes_mellitus 715 0.7796059
## 65          hepatic_failure 715 0.7796059
## 66          immunosuppression 715 0.7796059
## 67              leukemia 715 0.7796059
## 68              lymphoma 715 0.7796059
## 69      solid_tumor_with_metastasis 715 0.7796059
## 70              X 91713 100.0000000
```

## delete Column

reference for deleting attribute in columns with threshold 10%: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3701793/>

```
kolom_delete <- kolom_NA %>% filter(percen_freq > 10)
kolom_NA <- kolom_NA %>%
  filter(percen_freq <= 10)
data <- data %>% select(-c(kolom_delete$Nama_Kolom))
```

## Divide column by datatype with variable

```
# pembagian kolom NA berdasarkan tipe data
int_kat <- names(which(lapply(data[kolom_NA$Nama_Kolom],class) == "integer"))
flo_num <- names(which(lapply(data[kolom_NA$Nama_Kolom],class) == "numeric"))

kolom_NA_cat <- kolom_NA %>% filter(Nama_Kolom %in% int_kat) %>% select(Nama_Kolom)
kolom_NA_cat <- kolom_NA_cat$Nama_Kolom

kolom_NA_num <- kolom_NA %>% filter(Nama_Kolom %in% flo_num) %>% select(Nama_Kolom)
kolom_NA_num <- kolom_NA_num$Nama_Kolom
kolom_NA_num
```

```
## [1] "bmi" "height"
## [3] "weight" "apache_3j_diagnosis"
## [5] "resprate_apache" "temp_apache"
## [7] "d1_sysbp_noninvasive_min" "d1_temp_max"
## [9] "d1_temp_min" "apache_4a_hospital_death_prob"
## [11] "apache_4a_icu_death_prob"
```

```
kolom_NA_cat
```

```
## [1] "age" "apache_2_diagnosis"
## [3] "arf_apache" "gcs_eyes_apache"
## [5] "gcs_motor_apache" "gcs_unable_apache"
```

```
## [7] "gcs_verbal_apache"      "heart_rate_apache"
## [9] "intubated_apache"      "map_apache"
## [11] "ventilated_apache"     "d1_diasbp_max"
## [13] "d1_diasbp_min"         "d1_diasbp_noninvasive_max"
## [15] "d1_diasbp_noninvasive_min" "d1_heartrate_max"
## [17] "d1_heartrate_min"      "d1_mbp_max"
## [19] "d1_mbp_min"            "d1_mbp_noninvasive_max"
## [21] "d1_mbp_noninvasive_min" "d1_resprate_max"
## [23] "d1_resprate_min"       "d1_spo2_max"
## [25] "d1_spo2_min"           "d1_sysbp_max"
## [27] "d1_sysbp_min"          "d1_sysbp_noninvasive_max"
## [29] "h1_diasbp_max"         "h1_diasbp_min"
## [31] "h1_diasbp_noninvasive_max" "h1_diasbp_noninvasive_min"
## [33] "h1_heartrate_max"      "h1_heartrate_min"
## [35] "h1_mbp_max"            "h1_mbp_min"
## [37] "h1_mbp_noninvasive_max" "h1_mbp_noninvasive_min"
## [39] "h1_resprate_max"       "h1_resprate_min"
## [41] "h1_spo2_max"           "h1_spo2_min"
## [43] "h1_sysbp_max"          "h1_sysbp_min"
## [45] "h1_sysbp_noninvasive_max" "h1_sysbp_noninvasive_min"
## [47] "d1_glucose_max"        "d1_glucose_min"
## [49] "aids"                  "cirrhosis"
## [51] "diabetes_mellitus"     "hepatic_failure"
## [53] "immunosuppression"     "leukemia"
## [55] "lymphoma"              "solid_tumor_with_metastasis"
```

## median numeric imputation

```
for(kolomku in kolom_NA_num){
  data[[kolomku]][is.na(data[[kolomku]])] = median(data[[kolomku]],na.rm=T)
}

kolom_NA_cat_factor <- c()
kolom_NA_cat_int <- c()
for(kolom in kolom_NA_cat){
  banyak_unique <- length(unique(data[[kolom]]))
  if(banyak_unique > 10){
    kolom_NA_cat_int <- c(kolom_NA_cat_int,kolom)
  } else{
    kolom_NA_cat_factor <- c(kolom_NA_cat_factor,kolom)
  }
}
kolom_NA_cat_factor
```

```
## [1] "arf_apache"      "gcs_eyes_apache"
## [3] "gcs_motor_apache" "gcs_unable_apache"
## [5] "gcs_verbal_apache" "intubated_apache"
## [7] "ventilated_apache" "aids"
## [9] "cirrhosis"       "diabetes_mellitus"
## [11] "hepatic_failure" "immunosuppression"
## [13] "leukemia"        "lymphoma"
## [15] "solid_tumor_with_metastasis"
```

kolom\_NA\_cat\_int

```
## [1] "age" "apache_2_diagnosis"
## [3] "heart_rate_apache" "map_apache"
## [5] "d1_diasbp_max" "d1_diasbp_min"
## [7] "d1_diasbp_noninvasive_max" "d1_diasbp_noninvasive_min"
## [9] "d1_heartrate_max" "d1_heartrate_min"
## [11] "d1_mbp_max" "d1_mbp_min"
## [13] "d1_mbp_noninvasive_max" "d1_mbp_noninvasive_min"
## [15] "d1_resprate_max" "d1_resprate_min"
## [17] "d1_spo2_max" "d1_spo2_min"
## [19] "d1_sysbp_max" "d1_sysbp_min"
## [21] "d1_sysbp_noninvasive_max" "h1_diasbp_max"
## [23] "h1_diasbp_min" "h1_diasbp_noninvasive_max"
## [25] "h1_diasbp_noninvasive_min" "h1_heartrate_max"
## [27] "h1_heartrate_min" "h1_mbp_max"
## [29] "h1_mbp_min" "h1_mbp_noninvasive_max"
## [31] "h1_mbp_noninvasive_min" "h1_resprate_max"
## [33] "h1_resprate_min" "h1_spo2_max"
## [35] "h1_spo2_min" "h1_sysbp_max"
## [37] "h1_sysbp_min" "h1_sysbp_noninvasive_max"
## [39] "h1_sysbp_noninvasive_min" "d1_glucose_max"
## [41] "d1_glucose_min"
```

median integer imputation

```
for(kolomku in kolom_NA_cat_int){
  data[[kolomku]][is.na(data[[kolomku]])] = median(data[[kolomku]],na.rm=T)
}
```

mode factor imputation

```
for(kolomku in kolom_NA_cat_factor){
  data[[kolomku]][is.na(data[[kolomku]])] = as.numeric(names(sort(-table(data[[kolomku]]))) [1])
  data[[kolomku]] <- as.factor(data[[kolomku]])
}
```

Selecting Feature (Deleting id) because dont depending for evaluation.

```
data <- data %>% select(-c(encounter_id,patient_id,
                           hospital_id,icu_id))
```

Factorizing Dataset

```

kol_char <- names(which(lapply(data,class)=="character"))

for(kol in kol_char){
  mode_impute <- names(sort(-table(data[[kol]])))[1]
  data[[kol]] <- replace(data[[kol]],data[[kol]]=="",mode_impute)
  data[[kol]] <- as.factor(data[[kol]])
}
head(data)

```

```

##   age      bmi elective_surgery ethnicity gender height
## 1  68 22.73000              0 Caucasian      M  180.3
## 2  77 27.42000              0 Caucasian      F  160.0
## 3  25 31.95000              0 Caucasian      F  172.7
## 4  81 22.64000              1 Caucasian      F  165.1
## 5  19 27.65465              0 Caucasian      M  188.0
## 6  67 27.56000              0 Caucasian      M  190.5
##
##           icu_admit_source icu_stay_type      icu_type pre_icu_los_days weight
## 1                Floor          admit      CTICU      0.541666667    73.9
## 2                Floor          admit Med-Surg ICU      0.927777778    70.2
## 3      Accident & Emergency          admit Med-Surg ICU      0.000694444    95.3
## 4 Operating Room / Recovery          admit      CTICU      0.000694444    61.7
## 5      Accident & Emergency          admit Med-Surg ICU      0.073611111    80.3
## 6      Accident & Emergency          admit Med-Surg ICU      0.000694444   100.0
##
## apache_2_diagnosis apache_3j_diagnosis apache_post_operative arf_apache
## 1                113                502.01                0                0
## 2                108                203.01                0                0
## 3                122                703.03                0                0
## 4                203               1206.03                1                0
## 5                119                601.01                0                0
## 6                301                403.01                0                0
##
## gcs_eyes_apache gcs_motor_apache gcs_unable_apache gcs_verbal_apache
## 1                3                6                0                4
## 2                1                3                0                1
## 3                3                6                0                5
## 4                4                6                0                5
## 5                4                6                0                5
## 6                4                6                0                5
##
## heart_rate_apache intubated_apache map_apache resprate_apache temp_apache
## 1                118                0            40            36            39.3
## 2                120                0            46            33            35.1
## 3                102                0            68            37            36.7
## 4                114                1            60             4            34.8
## 5                 60                0           103            16            36.7
## 6                113                0           130            35            36.6
##
## ventilated_apache d1_diasbp_max d1_diasbp_min d1_diasbp_noninvasive_max
## 1                0                68            37                68
## 2                1                95            31                95
## 3                0                88            48                88
## 4                1                48            42                48
## 5                0                99            57                99
## 6                0               100            61               100
##
## d1_diasbp_noninvasive_min d1_heartrate_max d1_heartrate_min d1_mbp_max
## 1                37                119                72                89

```

## 2	31	118	72	120	
## 3	48	96	68	102	
## 4	42	116	92	84	
## 5	57	89	60	104	
## 6	61	113	83	127	
##	d1_mbp_min	d1_mbp_noninvasive_max	d1_mbp_noninvasive_min	d1_resprate_max	
## 1	46	89	46	34	
## 2	38	120	38	32	
## 3	68	102	68	21	
## 4	84	84	84	23	
## 5	90	104	90	18	
## 6	80	127	80	32	
##	d1_resprate_min	d1_spo2_max	d1_spo2_min	d1_sysbp_max	d1_sysbp_min
## 1	10	100	74	131	73
## 2	12	100	70	159	67
## 3	8	98	91	148	105
## 4	7	100	95	158	84
## 5	16	100	96	147	120
## 6	10	97	91	173	107
##	d1_sysbp_noninvasive_max	d1_sysbp_noninvasive_min	d1_temp_max	d1_temp_min	
## 1	131	73	39.9	37.2	
## 2	159	67	36.3	35.1	
## 3	148	105	37.0	36.7	
## 4	158	84	38.0	34.8	
## 5	147	120	37.2	36.7	
## 6	173	107	36.8	36.6	
##	h1_diasbp_max	h1_diasbp_min	h1_diasbp_noninvasive_max		
## 1	68	63	68		
## 2	61	48	61		
## 3	88	58	88		
## 4	62	44	74		
## 5	99	68	99		
## 6	89	89	89		
##	h1_diasbp_noninvasive_min	h1_heartrate_max	h1_heartrate_min	h1_mbp_max	
## 1	63	119	108	86	
## 2	48	114	100	85	
## 3	58	96	78	91	
## 4	62	100	96	92	
## 5	68	89	76	104	
## 6	89	83	83	111	
##	h1_mbp_min	h1_mbp_noninvasive_max	h1_mbp_noninvasive_min	h1_resprate_max	
## 1	85	86	85	26	
## 2	57	85	57	31	
## 3	83	91	83	20	
## 4	71	90	79	12	
## 5	92	104	92	21	
## 6	111	111	111	12	
##	h1_resprate_min	h1_spo2_max	h1_spo2_min	h1_sysbp_max	h1_sysbp_min
## 1	18	100	74	131	115
## 2	28	95	70	95	71
## 3	16	98	91	148	124
## 4	11	100	99	136	106
## 5	16	100	100	130	120
## 6	12	97	97	143	143



```
## h1_sysbp_noninvasive_max h1_sysbp_noninvasive_min d1_glucose_max
## 1 131 115 168
## 2 95 71 145
## 3 148 124 150
## 4 130 115 185
## 5 130 120 150
## 6 143 143 156
## d1_glucose_min apache_4a_hospital_death_prob apache_4a_icu_death_prob aids
## 1 109 0.10 0.05 0
## 2 128 0.47 0.29 0
## 3 107 0.00 0.00 0
## 4 88 0.04 0.03 0
## 5 107 0.05 0.02 0
## 6 125 0.05 0.02 0
## cirrhosis diabetes_mellitus hepatic_failure immunosuppression leukemia
## 1 0 1 0 0 0
## 2 0 1 0 0 0
## 3 0 0 0 0 0
## 4 0 0 0 0 0
## 5 0 0 0 0 0
## 6 0 1 0 0 0
## lymphoma solid_tumor_with_metastasis apache_3j_bodysystem apache_2_bodysystem
## 1 0 0 Sepsis Cardiovascular
## 2 0 0 Respiratory Respiratory
## 3 0 0 Metabolic Metabolic
## 4 0 0 Cardiovascular Cardiovascular
## 5 0 0 Trauma Trauma
## 6 0 0 Neurological Neurologic
## hospital_death
## 1 Survived
## 2 Survived
## 3 Survived
## 4 Survived
## 5 Survived
## 6 Survived
```

```
# data <- data %>%
# mutate(hospital_death = case_when(hospital_death == 0 ~ 'Survived',
#                                   hospital_death == 1 ~ 'Death'))
# data$hospital_death <- factor(data$hospital_death, levels = c("Survived", "Death"), labels = c("Survived", "Death"))
# str(data)
```

## feature engineering

### averaging dataset

```
data$d1_diasbp <- (data$d1_diasbp_min +
                  data$d1_diasbp_max)/2
data$d1_diasbp_noninvasive <- (data$d1_diasbp_noninvasive_min +
                              data$d1_diasbp_noninvasive_max)/2
data$d1_hearttrate <- (data$d1_hearttrate_min +
                      data$d1_hearttrate_max) / 2
```

```

data$d1_mbp <- (data$d1_mbp_max + data$d1_mbp_min)/2
data$d1_mbp_noninvasive <- (data$d1_mbp_noninvasive_max +
                             data$d1_mbp_noninvasive_min)/2
data$d1_glucose <- (data$d1_glucose_max + data$d1_glucose_min)/2
data$d1_resprate <- (data$d1_resprate_min + data$d1_resprate_max)/2
data$d1_sysbp <- (data$d1_sysbp_min + data$d1_sysbp_max)/2
data$d1_sysbp_noninvasive <- (data$d1_sysbp_noninvasive_max +
                               data$d1_sysbp_noninvasive_min)/2
data$d1_temp <- (data$d1_temp_min + data$d1_temp_max)/2
data$h1_diasbp <- (data$h1_diasbp_max + data$h1_diasbp_min)/2
data$h1_diasbp_noninvasive <- (data$h1_diasbp_noninvasive_min +
                                data$h1_diasbp_noninvasive_max)/2
data$h1_hearttrate <- (data$h1_hearttrate_min + data$h1_hearttrate_max)/2
data$h1_mbp <- (data$h1_mbp_max + data$h1_mbp_min)/2
data$h1_mbp_noninvasive <- (data$h1_mbp_noninvasive_max +
                             data$h1_mbp_noninvasive_min)/2
data$h1_resprate <- (data$h1_resprate_min + data$h1_resprate_max)
data$h1_spo2 <- (data$h1_spo2_min + data$h1_spo2_max)/2
data$h1_sysbp <- (data$h1_sysbp_max + data$h1_sysbp_min)/2
data$h1_sysbp_noninvasive <- (data$h1_sysbp_noninvasive_max +
                               data$h1_diasbp_noninvasive_min)/2

data <- data %>% select(-c(d1_diasbp_min,d1_diasbp_max,
                           d1_diasbp_noninvasive_max,
                           d1_diasbp_noninvasive_min,
                           d1_glucose_max,d1_glucose_min,
                           d1_hearttrate_max,d1_hearttrate_min,
                           d1_mbp_max,d1_mbp_min,
                           d1_mbp_noninvasive_max,
                           d1_mbp_min,d1_resprate_min,
                           d1_resprate_max,
                           d1_spo2_max,d1_spo2_min,
                           d1_sysbp_max,d1_sysbp_min,
                           d1_sysbp_max,d1_sysbp_min,
                           d1_sysbp_noninvasive_max,
                           d1_sysbp_noninvasive_min,
                           d1_temp_max,d1_temp_min, h1_diasbp_max,
                           h1_diasbp_min,h1_diasbp_noninvasive_max,
                           h1_diasbp_noninvasive_min,
                           h1_hearttrate_max,h1_hearttrate_min,
                           h1_mbp_min,h1_mbp_max,h1_mbp_noninvasive_max,
                           h1_mbp_noninvasive_min,
                           h1_resprate_max,h1_resprate_min,
                           h1_spo2_max,h1_spo2_min,
                           h1_sysbp_max,h1_sysbp_min,
                           h1_sysbp_noninvasive_min,
                           h1_sysbp_noninvasive_max))

names(data)

```

```

## [1] "age"                                "bmi"
## [3] "elective_surgery"                "ethnicity"
## [5] "gender"                          "height"

```

```
## [7] "icu_admit_source"      "icu_stay_type"
## [9] "icu_type"             "pre_icu_los_days"
## [11] "weight"               "apache_2_diagnosis"
## [13] "apache_3j_diagnosis"  "apache_post_operative"
## [15] "arf_apache"           "gcs_eyes_apache"
## [17] "gcs_motor_apache"     "gcs_unable_apache"
## [19] "gcs_verbal_apache"    "heart_rate_apache"
## [21] "intubated_apache"     "map_apache"
## [23] "resprate_apache"      "temp_apache"
## [25] "ventilated_apache"    "d1_mbp_noninvasive_min"
## [27] "apache_4a_hospital_death_prob" "apache_4a_icu_death_prob"
## [29] "aids"                 "cirrhosis"
## [31] "diabetes_mellitus"    "hepatic_failure"
## [33] "immunosuppression"    "leukemia"
## [35] "lymphoma"             "solid_tumor_with_metastasis"
## [37] "apache_3j_bodysystem" "apache_2_bodysystem"
## [39] "hospital_death"       "d1_diasbp"
## [41] "d1_diasbp_noninvasive" "d1_heartrate"
## [43] "d1_mbp"               "d1_mbp_noninvasive"
## [45] "d1_glucose"           "d1_resprate"
## [47] "d1_sysbp"             "d1_sysbp_noninvasive"
## [49] "d1_temp"              "h1_diasbp"
## [51] "h1_diasbp_noninvasive" "h1_heartrate"
## [53] "h1_mbp"               "h1_mbp_noninvasive"
## [55] "h1_resprate"          "h1_spo2"
## [57] "h1_sysbp"             "h1_sysbp_noninvasive"
```

## Delete Multicollinearity

singularities

```
model_logit <- glm(hospital_death ~ ., family=binomial, data=data)
summary(model_logit)
```

```
##
## Call:
## glm(formula = hospital_death ~ ., family = binomial, data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6246  -0.3656  -0.2248  -0.1353   3.5881
##
## Coefficients: (6 not defined because of singularities)
##              Estimate Std. Error z value
## (Intercept)    1.295e+01  9.317e-01  13.901
## age            2.447e-02  1.076e-03  22.735
## bmi           -2.323e-04  7.140e-03  -0.033
## elective_surgery -6.882e-01  9.434e-02 -7.295
## ethnicityAsian   2.722e-02  1.358e-01  0.201
## ethnicityCaucasian 5.955e-02  4.864e-02  1.224
## ethnicityHispanic 1.173e-01  8.029e-02  1.461
## ethnicityNative American 9.157e-02  1.549e-01  0.591
```

## ethnicityOther/Unknown	3.449e-02	8.041e-02	0.429
## genderM	8.122e-02	3.782e-02	2.147
## height	1.282e-03	3.146e-03	0.408
## icu_admit_sourceFloor	1.667e-01	3.636e-02	4.584
## icu_admit_sourceOperating Room / Recovery	-3.510e-01	1.010e-01	-3.475
## icu_admit_sourceOther Hospital	4.246e-01	7.359e-02	5.770
## icu_admit_sourceOther ICU	1.289e+00	1.266e-01	10.183
## icu_stay_typereadmit	2.300e-01	1.832e-01	1.255
## icu_stay_typetransfer	-2.968e-01	5.889e-02	-5.039
## icu_typeCCU-CTICU	-1.218e-01	7.687e-02	-1.585
## icu_typeCSICU	-4.809e-01	9.624e-02	-4.997
## icu_typeCTICU	-1.335e-01	9.771e-02	-1.366
## icu_typeMed-Surg ICU	-1.742e-01	6.030e-02	-2.889
## icu_typeMICU	-7.728e-02	7.068e-02	-1.093
## icu_typeNeuro ICU	2.799e-02	7.925e-02	0.353
## icu_typeSICU	-3.293e-02	8.356e-02	-0.394
## pre_icu_los_days	3.329e-02	4.356e-03	7.641
## weight	-2.016e-03	2.563e-03	-0.787
## apache_2_diagnosis	-1.362e-04	2.180e-04	-0.625
## apache_3j_diagnosis	3.202e-03	3.765e-04	8.505
## apache_post_operative	-3.298e+00	4.238e-01	-7.782
## arf_apache1	3.657e-01	7.265e-02	5.034
## gcs_eyes_apache2	-3.063e-01	6.598e-02	-4.641
## gcs_eyes_apache3	-3.881e-01	7.004e-02	-5.541
## gcs_eyes_apache4	-4.396e-01	7.440e-02	-5.908
## gcs_motor_apache2	2.241e-01	1.469e-01	1.526
## gcs_motor_apache3	2.124e-02	1.209e-01	0.176
## gcs_motor_apache4	-4.804e-01	6.538e-02	-7.347
## gcs_motor_apache5	-5.723e-01	7.013e-02	-8.160
## gcs_motor_apache6	-7.572e-01	8.019e-02	-9.443
## gcs_unable_apache1	9.200e-01	1.013e-01	9.086
## gcs_verbal_apache2	3.147e-01	8.243e-02	3.818
## gcs_verbal_apache3	1.043e-03	7.393e-02	0.014
## gcs_verbal_apache4	2.312e-01	6.258e-02	3.694
## gcs_verbal_apache5	-1.261e-01	6.320e-02	-1.995
## heart_rate_apache	1.944e-03	7.113e-04	2.733
## intubated_apache1	1.253e-02	3.843e-02	0.326
## map_apache	1.254e-03	3.659e-04	3.426
## resprate_apache	7.753e-03	1.257e-03	6.166
## temp_apache	-5.689e-02	2.027e-02	-2.806
## ventilated_apache1	8.658e-01	3.810e-02	22.726
## d1_mbp_noninvasive_min	-2.295e-02	1.285e-03	-17.862
## apache_4a_hospital_death_prob	9.991e-01	1.352e-01	7.392
## apache_4a_icu_death_prob	3.535e-01	1.485e-01	2.380
## aids1	3.868e-01	3.853e-01	1.004
## cirrhosis1	4.263e-01	1.047e-01	4.072
## diabetes_mellitus1	-2.274e-01	3.679e-02	-6.180
## hepatic_failure1	4.060e-01	1.127e-01	3.603
## immunosuppression1	3.021e-01	7.225e-02	4.181
## leukemia1	3.407e-01	1.240e-01	2.748
## lymphoma1	2.861e-01	1.632e-01	1.753
## solid_tumor_with_metastasis1	7.528e-01	7.641e-02	9.853
## apache_3j_bodysystemGastrointestinal	-6.454e-01	9.071e-02	-7.115
## apache_3j_bodysystemGenitourinary	-1.077e+01	6.318e+01	-0.170

## apache_3j_bodysystemGynecological	-1.166e+01	6.318e+01	-0.185
## apache_3j_bodysystemHematological	-1.949e+00	3.025e-01	-6.443
## apache_3j_bodysystemMetabolic	-3.179e+00	2.445e-01	-13.004
## apache_3j_bodysystemMusculoskeletal/Skin	-7.014e-01	3.597e-01	-1.950
## apache_3j_bodysystemNeurological	-7.963e-01	1.210e-01	-6.584
## apache_3j_bodysystemRespiratory	-4.038e-01	5.592e-02	-7.220
## apache_3j_bodysystemSepsis	-1.197e+00	1.496e-01	-7.998
## apache_3j_bodysystemTrauma	-1.511e+00	1.924e-01	-7.851
## apache_2_bodysystemGastrointestinal	NA	NA	NA
## apache_2_bodysystemHaematologic	NA	NA	NA
## apache_2_bodysystemMetabolic	NA	NA	NA
## apache_2_bodysystemNeurologic	NA	NA	NA
## apache_2_bodysystemRenal/Genitourinary	8.289e+00	6.318e+01	0.131
## apache_2_bodysystemRespiratory	NA	NA	NA
## apache_2_bodysystemTrauma	NA	NA	NA
## apache_2_bodysystemUndefined diagnoses	-1.426e+00	1.770e-01	-8.055
## apache_2_bodysystemUndefined Diagnoses	-2.520e+00	3.803e-01	-6.628
## d1_diasbp	-2.918e-02	1.011e-02	-2.887
## d1_diasbp_noninvasive	2.123e-02	1.024e-02	2.073
## d1_heartrate	1.139e-02	1.456e-03	7.820
## d1_mbp	-4.896e-03	6.528e-03	-0.750
## d1_mbp_noninvasive	1.175e-02	6.778e-03	1.733
## d1_glucose	2.117e-03	2.695e-04	7.857
## d1_resprate	1.216e-02	2.996e-03	4.058
## d1_sysbp	-1.632e-02	8.040e-03	-2.030
## d1_sysbp_noninvasive	6.473e-03	8.068e-03	0.802
## d1_temp	-3.314e-01	3.021e-02	-10.971
## h1_diasbp	5.582e-03	5.003e-03	1.116
## h1_diasbp_noninvasive	-2.883e-05	5.389e-03	-0.005
## h1_heartrate	-1.132e-03	1.078e-03	-1.050
## h1_mbp	-3.872e-03	4.834e-03	-0.801
## h1_mbp_noninvasive	-6.318e-03	4.853e-03	-1.302
## h1_resprate	9.915e-03	1.378e-03	7.198
## h1_spo2	-2.255e-02	2.497e-03	-9.033
## h1_sysbp	-5.225e-04	2.043e-03	-0.256
## h1_sysbp_noninvasive	2.212e-03	3.347e-03	0.661
##	Pr(> z )		
## (Intercept)	< 2e-16 ***		
## age	< 2e-16 ***		
## bmi	0.974053		
## elective_surgery	2.98e-13 ***		
## ethnicityAsian	0.841081		
## ethnicityCaucasian	0.220838		
## ethnicityHispanic	0.143890		
## ethnicityNative American	0.554532		
## ethnicityOther/Unknown	0.667923		
## genderM	0.031759 *		
## height	0.683608		
## icu_admit_sourceFloor	4.57e-06 ***		
## icu_admit_sourceOperating Room / Recovery	0.000511 ***		
## icu_admit_sourceOther Hospital	7.93e-09 ***		
## icu_admit_sourceOther ICU	< 2e-16 ***		
## icu_stay_typereadmit	0.209345		
## icu_stay_typetransfer	4.68e-07 ***		

## icu_typeCCU-CTICU	0.113037
## icu_typeCSICU	5.81e-07 ***
## icu_typeCTICU	0.171952
## icu_typeMed-Surg ICU	0.003861 **
## icu_typeMICU	0.274224
## icu_typeNeuro ICU	0.723917
## icu_typeSICU	0.693502
## pre_icu_los_days	2.16e-14 ***
## weight	0.431456
## apache_2_diagnosis	0.532262
## apache_3j_diagnosis	< 2e-16 ***
## apache_post_operative	7.14e-15 ***
## arf_apache1	4.81e-07 ***
## gcs_eyes_apache2	3.46e-06 ***
## gcs_eyes_apache3	3.00e-08 ***
## gcs_eyes_apache4	3.46e-09 ***
## gcs_motor_apache2	0.127041
## gcs_motor_apache3	0.860543
## gcs_motor_apache4	2.03e-13 ***
## gcs_motor_apache5	3.36e-16 ***
## gcs_motor_apache6	< 2e-16 ***
## gcs_unable_apache1	< 2e-16 ***
## gcs_verbal_apache2	0.000134 ***
## gcs_verbal_apache3	0.988749
## gcs_verbal_apache4	0.000221 ***
## gcs_verbal_apache5	0.046074 *
## heart_rate_apache	0.006273 **
## intubated_apache1	0.744467
## map_apache	0.000612 ***
## resprate_apache	7.01e-10 ***
## temp_apache	0.005015 **
## ventilated_apache1	< 2e-16 ***
## dl_mbp_noninvasive_min	< 2e-16 ***
## apache_4a_hospital_death_prob	1.44e-13 ***
## apache_4a_icu_death_prob	0.017316 *
## aids1	0.315349
## cirrhosis1	4.66e-05 ***
## diabetes_mellitus1	6.40e-10 ***
## hepatic_failure1	0.000315 ***
## immunosuppression1	2.90e-05 ***
## leukemia1	0.006002 **
## lymphoma1	0.079522 .
## solid_tumor_with_metastasis1	< 2e-16 ***
## apache_3j_bodysystemGastrointestinal	1.12e-12 ***
## apache_3j_bodysystemGenitourinary	0.864702
## apache_3j_bodysystemGynecological	0.853569
## apache_3j_bodysystemHematological	1.17e-10 ***
## apache_3j_bodysystemMetabolic	< 2e-16 ***
## apache_3j_bodysystemMusculoskeletal/Skin	0.051190 .
## apache_3j_bodysystemNeurological	4.59e-11 ***
## apache_3j_bodysystemRespiratory	5.19e-13 ***
## apache_3j_bodysystemSepsis	1.26e-15 ***
## apache_3j_bodysystemTrauma	4.13e-15 ***
## apache_2_bodysystemGastrointestinal	NA

```
## apache_2_bodysystemHaematologic      NA
## apache_2_bodysystemMetabolic          NA
## apache_2_bodysystemNeurologic         NA
## apache_2_bodysystemRenal/Genitourinary 0.895620
## apache_2_bodysystemRespiratory        NA
## apache_2_bodysystemTrauma             NA
## apache_2_bodysystemUndefined diagnoses 7.95e-16 ***
## apache_2_bodysystemUndefined Diagnoses 3.40e-11 ***
## d1_diasbp                             0.003887 **
## d1_diasbp_noninvasive                  0.038132 *
## d1_heartrate                          5.28e-15 ***
## d1_mbp                                0.453262
## d1_mbp_noninvasive                    0.083056 .
## d1_glucose                            3.94e-15 ***
## d1_resprate                           4.94e-05 ***
## d1_sysbp                              0.042369 *
## d1_sysbp_noninvasive                  0.422388
## d1_temp                               < 2e-16 ***
## h1_diasbp                             0.264452
## h1_diasbp_noninvasive                 0.995732
## h1_heartrate                         0.293619
## h1_mbp                                0.423143
## h1_mbp_noninvasive                   0.192945
## h1_resprate                          6.13e-13 ***
## h1_spo2                               < 2e-16 ***
## h1_sysbp                              0.798175
## h1_sysbp_noninvasive                 0.508757
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 53908  on 91712  degrees of freedom
## Residual deviance: 38628  on 91621  degrees of freedom
## AIC: 38812
##
## Number of Fisher Scoring iterations: 11
```

ada attribute yang bernilai NA dikarenakan ada suatu Singularitas, atau kemungkinan ada suatu Multicollinearity

**Dikarenakan adanya singularitas, maka ada Multicollinearity**

```
data <- data %>% select(-c apache_2_bodysystem))
model_logit <- glm(hospital_death ~ ., family=binomial, data=data)
summary(model_logit)
```

```
##
## Call:
## glm(formula = hospital_death ~ ., family = binomial, data = data)
##
## Deviance Residuals:
```

```

##      Min      1Q   Median      3Q      Max
## -2.6217  -0.3663  -0.2267  -0.1376   3.5943
##
## Coefficients:
##
##              Estimate Std. Error z value
## (Intercept)    13.2044509   0.9317495   14.172
## age             0.0243373   0.0010769   22.598
## bmi             0.0003327   0.0071374    0.047
## elective_surgery -0.7477388   0.0937334   -7.977
## ethnicityAsian   0.0246090   0.1356087    0.181
## ethnicityCaucasian 0.0564878   0.0486330    1.162
## ethnicityHispanic 0.1167822   0.0802230    1.456
## ethnicityNative American 0.0896774   0.1548415    0.579
## ethnicityOther/Unknown 0.0303726   0.0803523    0.378
## genderM         0.0739105   0.0378289    1.954
## height          0.0013751   0.0031460    0.437
## icu_admit_sourceFloor 0.1709878   0.0363851    4.699
## icu_admit_sourceOperating Room / Recovery -0.3423652   0.1012424   -3.382
## icu_admit_sourceOther Hospital 0.4299846   0.0736350    5.839
## icu_admit_sourceOther ICU 1.3870554   0.1259131   11.016
## icu_stay_typereadmit 0.2650091   0.1828207    1.450
## icu_stay_typetransfer -0.2913183   0.0587669   -4.957
## icu_typeCCU-CTICU -0.1328844   0.0768388   -1.729
## icu_typeCSICU    -0.5148517   0.0960569   -5.360
## icu_typeCTICU    -0.1927730   0.0973279   -1.981
## icu_typeMed-Surg ICU -0.1737967   0.0603245   -2.881
## icu_typeMICU     -0.0784274   0.0707182   -1.109
## icu_typeNeuro ICU  0.0297931   0.0793354    0.376
## icu_typeSICU     -0.0429090   0.0835262   -0.514
## pre_icu_los_days  0.0328321   0.0043686    7.515
## weight          -0.0022352   0.0025624   -0.872
## apache_2_diagnosis -0.0005280   0.0002161   -2.443
## apache_3j_diagnosis 0.0026771   0.0003682    7.271
## apache_post_operative -2.7715716   0.4125770   -6.718
## arf_apache1      0.3708484   0.0726017    5.108
## gcs_eyes_apache2 -0.3011241   0.0660114   -4.562
## gcs_eyes_apache3 -0.3859150   0.0700414   -5.510
## gcs_eyes_apache4 -0.4343039   0.0743959   -5.838
## gcs_motor_apache2 0.2387984   0.1470905    1.623
## gcs_motor_apache3 0.0333695   0.1209099    0.276
## gcs_motor_apache4 -0.4650679   0.0653510   -7.116
## gcs_motor_apache5 -0.5531070   0.0701040   -7.890
## gcs_motor_apache6 -0.7409417   0.0801029   -9.250
## gcs_unable_apache1 0.9158402   0.1010952    9.059
## gcs_verbal_apache2 0.3085180   0.0824544    3.742
## gcs_verbal_apache3 0.0034636   0.0738690    0.047
## gcs_verbal_apache4 0.2324079   0.0625426    3.716
## gcs_verbal_apache5 -0.1162570   0.0630917   -1.843
## heart_rate_apache 0.0018961   0.0007115    2.665
## intubated_apache1 0.0009896   0.0383582    0.026
## map_apache       0.0012865   0.0003651    3.523
## resprate_apache  0.0076883   0.0012554    6.124
## temp_apache      -0.0509630   0.0202678   -2.514
## ventilated_apache1 0.8464675   0.0380977   22.218

```



## d1_mbp_noninvasive_min	-0.0231848	0.0012848	-18.045
## apache_4a_hospital_death_prob	1.0829532	0.1360271	7.961
## apache_4a_icu_death_prob	0.3286380	0.1494344	2.199
## aids1	0.3928768	0.3849558	1.021
## cirrhosis1	0.4273153	0.1046700	4.082
## diabetes_mellitus1	-0.2395560	0.0366886	-6.529
## hepatic_failure1	0.3994222	0.1126018	3.547
## immunosuppression1	0.3058148	0.0723017	4.230
## leukemia1	0.3347209	0.1240654	2.698
## lymphoma1	0.2849948	0.1632175	1.746
## solid_tumor_with_metastasis1	0.7562412	0.0765511	9.879
## apache_3j_bodysystemGastrointestinal	-0.4662466	0.0893951	-5.216
## apache_3j_bodysystemGenitourinary	-1.9870340	0.2978638	-6.671
## apache_3j_bodysystemGynecological	-2.8734089	0.7672454	-3.745
## apache_3j_bodysystemHematological	-1.5059689	0.2967752	-5.074
## apache_3j_bodysystemMetabolic	-2.8167111	0.2392724	-11.772
## apache_3j_bodysystemMusculoskeletal/Skin	-1.9894162	0.3455704	-5.757
## apache_3j_bodysystemNeurological	-0.5796708	0.1184559	-4.894
## apache_3j_bodysystemRespiratory	-0.3196892	0.0555153	-5.759
## apache_3j_bodysystemSepsis	-0.9804454	0.1467636	-6.680
## apache_3j_bodysystemTrauma	-1.2227654	0.1894552	-6.454
## d1_diasbp	-0.0291667	0.0100884	-2.891
## d1_diasbp_noninvasive	0.0209703	0.0102222	2.051
## d1_heartrate	0.0113617	0.0014571	7.798
## d1_mbp	-0.0047922	0.0065151	-0.736
## d1_mbp_noninvasive	0.0119046	0.0067673	1.759
## d1_glucose	0.0021970	0.0002695	8.153
## d1_resprate	0.0116968	0.0029964	3.904
## d1_sysbp	-0.0168286	0.0080933	-2.079
## d1_sysbp_noninvasive	0.0071075	0.0081221	0.875
## d1_temp	-0.3421224	0.0302051	-11.327
## h1_diasbp	0.0071055	0.0049590	1.433
## h1_diasbp_noninvasive	-0.0013949	0.0053507	-0.261
## h1_heartrate	-0.0010860	0.0010781	-1.007
## h1_mbp	-0.0029331	0.0047865	-0.613
## h1_mbp_noninvasive	-0.0073600	0.0048064	-1.531
## h1_resprate	0.0102127	0.0013767	7.418
## h1_spo2	-0.0229028	0.0024881	-9.205
## h1_sysbp	-0.0005036	0.0020410	-0.247
## h1_sysbp_noninvasive	0.0022868	0.0033455	0.684
##	Pr(> z )		
## (Intercept)	< 2e-16 ***		
## age	< 2e-16 ***		
## bmi	0.962818		
## elective_surgery	1.50e-15 ***		
## ethnicityAsian	0.855998		
## ethnicityCaucasian	0.245434		
## ethnicityHispanic	0.145470		
## ethnicityNative American	0.562484		
## ethnicityOther/Unknown	0.705436		
## genderM	0.050724 .		
## height	0.662043		
## icu_admit_sourceFloor	2.61e-06 ***		
## icu_admit_sourceOperating Room / Recovery	0.000721 ***		

## icu_admit_sourceOther Hospital	5.24e-09 ***
## icu_admit_sourceOther ICU	< 2e-16 ***
## icu_stay_typereadmit	0.147182
## icu_stay_typetransfer	7.15e-07 ***
## icu_typeCCU-CTICU	0.083739 .
## icu_typeCSICU	8.33e-08 ***
## icu_typeCTICU	0.047630 *
## icu_typeMed-Surg ICU	0.003964 **
## icu_typeMICU	0.267425
## icu_typeNeuro ICU	0.707264
## icu_typeSICU	0.607448
## pre_icu_los_days	5.67e-14 ***
## weight	0.383037
## apache_2_diagnosis	0.014546 *
## apache_3j_diagnosis	3.57e-13 ***
## apache_post_operative	1.85e-11 ***
## arf_apache1	3.26e-07 ***
## gcs_eyes_apache2	5.07e-06 ***
## gcs_eyes_apache3	3.59e-08 ***
## gcs_eyes_apache4	5.29e-09 ***
## gcs_motor_apache2	0.104487
## gcs_motor_apache3	0.782559
## gcs_motor_apache4	1.11e-12 ***
## gcs_motor_apache5	3.03e-15 ***
## gcs_motor_apache6	< 2e-16 ***
## gcs_unable_apache1	< 2e-16 ***
## gcs_verbal_apache2	0.000183 ***
## gcs_verbal_apache3	0.962602
## gcs_verbal_apache4	0.000202 ***
## gcs_verbal_apache5	0.065378 .
## heart_rate_apache	0.007704 **
## intubated_apache1	0.979418
## map_apache	0.000426 ***
## resprate_apache	9.11e-10 ***
## temp_apache	0.011921 *
## ventilated_apache1	< 2e-16 ***
## d1_mbp_noninvasive_min	< 2e-16 ***
## apache_4a_hospital_death_prob	1.70e-15 ***
## apache_4a_icu_death_prob	0.027863 *
## aids1	0.307455
## cirrhosis1	4.46e-05 ***
## diabetes_mellitus1	6.60e-11 ***
## hepatic_failure1	0.000389 ***
## immunosuppression1	2.34e-05 ***
## leukemia1	0.006977 **
## lymphoma1	0.080793 .
## solid_tumor_with_metastasis1	< 2e-16 ***
## apache_3j_bodysystemGastrointestinal	1.83e-07 ***
## apache_3j_bodysystemGenitourinary	2.54e-11 ***
## apache_3j_bodysystemGynecological	0.000180 ***
## apache_3j_bodysystemHematological	3.89e-07 ***
## apache_3j_bodysystemMetabolic	< 2e-16 ***
## apache_3j_bodysystemMusculoskeletal/Skin	8.57e-09 ***
## apache_3j_bodysystemNeurological	9.90e-07 ***

```
## apache_3j_bodysystemRespiratory      8.48e-09 ***
## apache_3j_bodysystemSepsis           2.38e-11 ***
## apache_3j_bodysystemTrauma           1.09e-10 ***
## d1_diasbp                            0.003839 **
## d1_diasbp_noninvasive                 0.040223 *
## d1_heartrate                         6.31e-15 ***
## d1_mbp                               0.461999
## d1_mbp_noninvasive                   0.078554 .
## d1_glucose                          3.54e-16 ***
## d1_resprate                         9.48e-05 ***
## d1_sysbp                            0.037587 *
## d1_sysbp_noninvasive                 0.381529
## d1_temp                             < 2e-16 ***
## h1_diasbp                           0.151904
## h1_diasbp_noninvasive                0.794333
## h1_heartrate                        0.313780
## h1_mbp                              0.540019
## h1_mbp_noninvasive                   0.125699
## h1_resprate                        1.19e-13 ***
## h1_spo2                             < 2e-16 ***
## h1_sysbp                           0.805120
## h1_sysbp_noninvasive                 0.494256
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 53908  on 91712  degrees of freedom
## Residual deviance: 38733  on 91624  degrees of freedom
## AIC: 38911
##
## Number of Fisher Scoring iterations: 7
```

Dihapus Atribut yang bernilai VIF lebih dari 5%

```
data <- data %>% select(-c(d1_sysbp_noninvasive,apache_3j_diagnosis,
                          d1_diasbp_noninvasive,d1_mbp_noninvasive,
                          h1_mbp_noninvasive,h1_diasbp_noninvasive))
model_logit <- glm(hospital_death ~ .,family=binomial,data=data)
summary(model_logit)

##
## Call:
## glm(formula = hospital_death ~ ., family = binomial, data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6203  -0.3667  -0.2277  -0.1391   3.7398
##
## Coefficients:
##
##              Estimate Std. Error z value
## (Intercept)    13.3981163  0.9316307  14.381
```

## age	0.0241325	0.0010728	22.495
## bmi	0.0003668	0.0071219	0.052
## elective_surgery	-0.7580741	0.0925732	-8.189
## ethnicityAsian	0.0160493	0.1354371	0.118
## ethnicityCaucasian	0.0540729	0.0486152	1.112
## ethnicityHispanic	0.1125326	0.0801777	1.404
## ethnicityNative American	0.0930028	0.1547638	0.601
## ethnicityOther/Unknown	0.0289549	0.0803077	0.361
## genderM	0.0738126	0.0378101	1.952
## height	0.0014122	0.0031407	0.450
## icu_admit_sourceFloor	0.1702969	0.0363767	4.681
## icu_admit_sourceOperating Room / Recovery	-0.3412431	0.1010300	-3.378
## icu_admit_sourceOther Hospital	0.4273718	0.0735508	5.811
## icu_admit_sourceOther ICU	1.6716167	0.1190664	14.039
## icu_stay_typereadmit	0.2814729	0.1828237	1.540
## icu_stay_typetransfer	-0.2538450	0.0579494	-4.380
## icu_typeCCU-CTICU	-0.1340366	0.0767460	-1.746
## icu_typeCSICU	-0.4981572	0.0959095	-5.194
## icu_typeCTICU	-0.1710235	0.0970516	-1.762
## icu_typeMed-Surg ICU	-0.1681338	0.0602101	-2.792
## icu_typeMICU	-0.0757945	0.0706353	-1.073
## icu_typeNeuro ICU	0.0431562	0.0792429	0.545
## icu_typeSICU	-0.0381199	0.0833585	-0.457
## pre_icu_los_days	0.0322615	0.0043788	7.368
## weight	-0.0022286	0.0025567	-0.872
## apache_2_diagnosis	-0.0004867	0.0002148	-2.266
## apache_post_operative	0.1191797	0.1151114	1.035
## arf_apache1	0.3639008	0.0725025	5.019
## gcs_eyes_apache2	-0.3006715	0.0660096	-4.555
## gcs_eyes_apache3	-0.3867796	0.0700390	-5.522
## gcs_eyes_apache4	-0.4331509	0.0743632	-5.825
## gcs_motor_apache2	0.2374621	0.1471861	1.613
## gcs_motor_apache3	0.0437038	0.1209618	0.361
## gcs_motor_apache4	-0.4593810	0.0653612	-7.028
## gcs_motor_apache5	-0.5464054	0.0701557	-7.788
## gcs_motor_apache6	-0.7369319	0.0801229	-9.198
## gcs_unable_apache1	0.8991899	0.1010340	8.900
## gcs_verbal_apache2	0.3107158	0.0824005	3.771
## gcs_verbal_apache3	0.0032354	0.0738672	0.044
## gcs_verbal_apache4	0.2335764	0.0625116	3.737
## gcs_verbal_apache5	-0.1036367	0.0630306	-1.644
## heart_rate_apache	0.0019380	0.0007118	2.723
## intubated_apache1	-0.0010243	0.0383681	-0.027
## map_apache	0.0012651	0.0003642	3.473
## resprate_apache	0.0076805	0.0012527	6.131
## temp_apache	-0.0486872	0.0202894	-2.400
## ventilated_apache1	0.8343804	0.0379863	21.965
## dl_mbp_noninvasive_min	-0.0226566	0.0012675	-17.876
## apache_4a_hospital_death_prob	1.0959254	0.1356675	8.078
## apache_4a_icu_death_prob	0.3352212	0.1491987	2.247
## aids1	0.3964544	0.3846255	1.031
## cirrhosis1	0.4154687	0.1044862	3.976
## diabetes_mellitus1	-0.2480992	0.0366322	-6.773
## hepatic_failure1	0.3992688	0.1122835	3.556

## immunosuppression1	0.3034573	0.0721981	4.203
## leukemia1	0.3233097	0.1239723	2.608
## lymphoma1	0.2813025	0.1630453	1.725
## solid_tumor_with_metastasis1	0.7458724	0.0763744	9.766
## apache_3j_bodysystemGastrointestinal	0.0504322	0.0554181	0.910
## apache_3j_bodysystemGenitourinary	0.0212062	0.1057222	0.201
## apache_3j_bodysystemGynecological	-1.2579784	0.7336437	-1.715
## apache_3j_bodysystemHematological	0.3180563	0.1589497	2.001
## apache_3j_bodysystemMetabolic	-1.2491257	0.1043241	-11.974
## apache_3j_bodysystemMusculoskeletal/Skin	0.2089520	0.1549879	1.348
## apache_3j_bodysystemNeurological	0.1803356	0.0567796	3.176
## apache_3j_bodysystemRespiratory	-0.0783944	0.0449349	-1.745
## apache_3j_bodysystemSepsis	0.0446141	0.0438163	1.018
## apache_3j_bodysystemTrauma	0.0300445	0.0801309	0.375
## d1_diasbp	-0.0078234	0.0024293	-3.220
## d1_heartrate	0.0113233	0.0014577	7.768
## d1_mbp	0.0056063	0.0026926	2.082
## d1_glucose	0.0022149	0.0002694	8.223
## d1_resprate	0.0115243	0.0029980	3.844
## d1_sysbp	-0.0095729	0.0014873	-6.436
## d1_temp	-0.3407017	0.0302804	-11.252
## h1_diasbp	0.0052778	0.0025246	2.091
## h1_heartrate	-0.0011068	0.0010766	-1.028
## h1_mbp	-0.0092862	0.0023455	-3.959
## h1_resprate	0.0101250	0.0013760	7.358
## h1_spo2	-0.0228517	0.0024842	-9.199
## h1_sysbp	-0.0004683	0.0018680	-0.251
## h1_sysbp_noninvasive	0.0020242	0.0028708	0.705
##	Pr(> z )		
## (Intercept)	< 2e-16 ***		
## age	< 2e-16 ***		
## bmi	0.958921		
## elective_surgery	2.64e-16 ***		
## ethnicityAsian	0.905672		
## ethnicityCaucasian	0.266025		
## ethnicityHispanic	0.160456		
## ethnicityNative American	0.547884		
## ethnicityOther/Unknown	0.718436		
## genderM	0.050915 .		
## height	0.652965		
## icu_admit_sourceFloor	2.85e-06 ***		
## icu_admit_sourceOperating Room / Recovery	0.000731 ***		
## icu_admit_sourceOther Hospital	6.23e-09 ***		
## icu_admit_sourceOther ICU	< 2e-16 ***		
## icu_stay_typereadmit	0.123661		
## icu_stay_typetransfer	1.18e-05 ***		
## icu_typeCCU-CTICU	0.080725 .		
## icu_typeCSICU	2.06e-07 ***		
## icu_typeCTICU	0.078037 .		
## icu_typeMed-Surg ICU	0.005231 **		
## icu_typeMICU	0.283253		
## icu_typeNeuro ICU	0.586024		
## icu_typeSICU	0.647455		
## pre_icu_los_days	1.74e-13 ***		

## weight	0.383381
## apache_2_diagnosis	0.023437 *
## apache_post_operative	0.300509
## arf_apache1	5.19e-07 ***
## gcs_eyes_apache2	5.24e-06 ***
## gcs_eyes_apache3	3.34e-08 ***
## gcs_eyes_apache4	5.72e-09 ***
## gcs_motor_apache2	0.106670
## gcs_motor_apache3	0.717873
## gcs_motor_apache4	2.09e-12 ***
## gcs_motor_apache5	6.78e-15 ***
## gcs_motor_apache6	< 2e-16 ***
## gcs_unable_apache1	< 2e-16 ***
## gcs_verbal_apache2	0.000163 ***
## gcs_verbal_apache3	0.965063
## gcs_verbal_apache4	0.000187 ***
## gcs_verbal_apache5	0.100129
## heart_rate_apache	0.006478 **
## intubated_apache1	0.978702
## map_apache	0.000514 ***
## resprate_apache	8.72e-10 ***
## temp_apache	0.016411 *
## ventilated_apache1	< 2e-16 ***
## d1_mbp_noninvasive_min	< 2e-16 ***
## apache_4a_hospital_death_prob	6.58e-16 ***
## apache_4a_icu_death_prob	0.024652 *
## aids1	0.302656
## cirrhosis1	7.00e-05 ***
## diabetes_mellitus1	1.26e-11 ***
## hepatic_failure1	0.000377 ***
## immunosuppression1	2.63e-05 ***
## leukemia1	0.009109 **
## lymphoma1	0.084473 .
## solid_tumor_with_metastasis1	< 2e-16 ***
## apache_3j_bodysystemGastrointestinal	0.362806
## apache_3j_bodysystemGenitourinary	0.841024
## apache_3j_bodysystemGynecological	0.086400 .
## apache_3j_bodysystemHematological	0.045394 *
## apache_3j_bodysystemMetabolic	< 2e-16 ***
## apache_3j_bodysystemMusculoskeletal/Skin	0.177600
## apache_3j_bodysystemNeurological	0.001493 **
## apache_3j_bodysystemRespiratory	0.081051 .
## apache_3j_bodysystemSepsis	0.308580
## apache_3j_bodysystemTrauma	0.707703
## d1_diasbp	0.001280 **
## d1_heartrate	7.98e-15 ***
## d1_mbp	0.037332 *
## d1_glucose	< 2e-16 ***
## d1_resprate	0.000121 ***
## d1_sysbp	1.22e-10 ***
## d1_temp	< 2e-16 ***
## h1_diasbp	0.036571 *
## h1_heartrate	0.303930
## h1_mbp	7.52e-05 ***

```
## h1_resprate                1.87e-13 ***
## h1_spo2                    < 2e-16 ***
## h1_sysbp                   0.802070
## h1_sysbp_noninvasive       0.480758
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 53908  on 91712  degrees of freedom
## Residual deviance: 38800  on 91630  degrees of freedom
## AIC: 38966
##
## Number of Fisher Scoring iterations: 7
```

```
df_vif <- vif(model_logit)
df_vif <- as.data.frame(df_vif)

colnames(df_vif) <- c("GVIF", "Df", "VIF")
df_vif_fil <- df_vif %>% filter(VIF>=5)
df_vif <- df_vif %>% arrange(VIF)
df_vif
```

	GVIF	Df	VIF
## aids	1.008823	1	1.004402
## lymphoma	1.018196	1	1.009057
## ethnicity	1.193287	5	1.017828
## leukemia	1.046124	1	1.022802
## arf_apache	1.052674	1	1.025999
## h1_spo2	1.054743	1	1.027007
## icu_type	1.855363	7	1.045138
## icu_stay_type	1.255271	2	1.058484
## apache_3j_bodysystem	3.917740	10	1.070660
## solid_tumor_with_metastasis	1.148545	1	1.071702
## gcs_unable_apache	1.153381	1	1.073956
## immunosuppression	1.180378	1	1.086452
## diabetes_mellitus	1.196329	1	1.093768
## d1_glucose	1.205784	1	1.098082
## pre_icu_los_days	1.239007	1	1.113107
## age	1.296938	1	1.138832
## map_apache	1.466491	1	1.210988
## gcs_motor_apache	7.879418	5	1.229276
## cirrhosis	1.526697	1	1.235596
## hepatic_failure	1.532908	1	1.238106
## gcs_verbal_apache	5.829909	4	1.246544
## intubated_apache	1.605045	1	1.266904
## apache_2_diagnosis	1.675275	1	1.294324
## icu_admit_source	8.397254	4	1.304720
## h1_resprate	1.881266	1	1.371593
## d1_mbp_noninvasive_min	1.888473	1	1.374217
## resprate_apache	1.912001	1	1.382751
## ventilated_apache	1.914615	1	1.383696
## gender	1.937537	1	1.391954
## gcs_eyes_apache	7.568812	3	1.401214

```
## d1_resprate      2.176024  1 1.475135
## temp_apache      2.740330  1 1.655394
## heart_rate_apache 2.745408  1 1.656927
## d1_temp          2.774842  1 1.665786
## h1_heartrate     2.835997  1 1.684042
## elective_surgery  3.618180  1 1.902152
## d1_heartrate     3.691996  1 1.921457
## d1_sysbp         3.949325  1 1.987291
## d1_diasbp        4.973844  1 2.230212
## height           6.192468  1 2.488467
## apache_4a_icu_death_prob 6.582243  1 2.565588
## apache_4a_hospital_death_prob 7.047798  1 2.654769
## apache_post_operative 7.132021  1 2.670584
## d1_mbp           7.583670  1 2.753846
## h1_diasbp        8.397735  1 2.897885
## h1_mbp           9.399459  1 3.065854
## h1_sysbp        11.528199  1 3.395320
## h1_sysbp_noninvasive 15.617315  1 3.951875
## bmi             18.243810  1 4.271277
## weight           21.303482  1 4.615570
```

## Check data after Feature Selection

```
summary(data)
```

```
##          age          bmi      elective_surgery      ethnicity
##  Min.   :16.00  Min.   :14.84  Min.   :0.0000  African American: 9547
##  1st Qu.:53.00  1st Qu.:23.79  1st Qu.:0.0000  Asian           : 1129
##  Median :65.00  Median :27.65  Median :0.0000  Caucasian       :72079
##  Mean   :62.43  Mean   :29.13  Mean   :0.1837  Hispanic        : 3796
##  3rd Qu.:75.00  3rd Qu.:32.65  3rd Qu.:0.0000  Native American :  788
##  Max.   :89.00  Max.   :67.81  Max.   :1.0000  Other/Unknown   : 4374
##
##  gender      height      icu_admit_source  icu_stay_type
##  F:42219  Min.   :137.2  Accident & Emergency  :54172  admit   :86183
##  M:49494  1st Qu.:162.6  Floor                 :15611  readmit :  560
##           Median :170.1  Operating Room / Recovery:18713  transfer: 4970
##           Mean   :169.6  Other Hospital         : 2358
##           3rd Qu.:177.8  Other ICU              :  859
##           Max.   :195.6
##
##          icu_type      pre_icu_los_days      weight      apache_2_diagnosis
##  Med-Surg ICU:50586  Min.   : -24.94722  Min.   : 38.60  Min.   :101.0
##  MICU              : 7695  1st Qu.:  0.03542  1st Qu.: 67.30  1st Qu.:113.0
##  Neuro ICU         : 7675  Median :  0.13889  Median : 80.30  Median :122.0
##  CCU-CTICU         : 7156  Mean   :  0.83577  Mean   : 83.92  Mean   :184.3
##  SICU              : 5209  3rd Qu.:  0.40903  3rd Qu.: 96.30  3rd Qu.:301.0
##  Cardiac ICU       : 4776  Max.   :159.09097  Max.   :186.00  Max.   :308.0
##  (Other)           : 8616
##  apache_post_operative  arf_apache  gcs_eyes_apache  gcs_motor_apache
##  Min.   :0.0000      0:89167      1: 8274      1: 5543
##  1st Qu.:0.0000      1: 2546      2: 4680      2:  309
```



```

## Median :0.0000          3:13863          3: 524
## Mean   :0.2011          4:64896          4: 4494
## 3rd Qu.:0.0000          5: 7982
## Max.   :1.0000          6:72861
##
## gcs_unable_apache gcs_verbal_apache heart_rate_apache intubated_apache
## 0:90849          1:16741          Min.   : 30.00    0:77952
## 1: 864           2: 1940          1st Qu.: 87.00    1:13761
##                3: 3275          Median :104.00
##                4:10947          Mean   : 99.75
##                5:58810          3rd Qu.:120.00
##                Max.   :178.00
##
## map_apache      resprate_apache temp_apache      ventilated_apache
## Min.   : 40.00   Min.   : 4.00   Min.   :32.10   0:62073
## 1st Qu.: 54.00   1st Qu.:11.00   1st Qu.:36.20   1:29640
## Median : 67.00   Median :28.00   Median :36.50
## Mean   : 87.79   Mean   :25.84   Mean   :36.42
## 3rd Qu.:124.00   3rd Qu.:36.00   3rd Qu.:36.70
## Max.   :200.00   Max.   :60.00   Max.   :39.70
##
## d1_mbp_noninvasive_min apache_4a_hospital_death_prob apache_4a_icu_death_prob
## Min.   : 22.00     Min.   : -1.0000     Min.   : -1.00000
## 1st Qu.: 55.00     1st Qu.: 0.0200     1st Qu.: 0.01000
## Median : 64.00     Median : 0.0500     Median : 0.02000
## Mean   : 64.93     Mean   : 0.0836     Mean   : 0.04188
## 3rd Qu.: 74.00     3rd Qu.: 0.1200     3rd Qu.: 0.06000
## Max.   :112.00     Max.   : 0.9900     Max.   : 0.97000
##
## aids      cirrhosis diabetes_mellitus hepatic_failure immunosuppression
## 0:91635    0:90285    0:71221          0:90531          0:89332
## 1: 78      1: 1428    1:20492          1: 1182          1: 2381
##
##
##
##
## leukemia lymphoma solid_tumor_with_metastasis      apache_3j_bodysystem
## 0:91070    0:91337    0:89835          Cardiovascular :31661
## 1: 643     1: 376     1: 1878          Neurological   :11896
##                Sepsis           :11740
##                Respiratory      :11609
##                Gastrointestinal: 9026
##                Metabolic        : 7650
##                (Other)         : 8131
##
## hospital_death d1_diasbp      d1_heartrate      d1_mbp
## Survived:83798 Min.   : 29.50   Min.   : 29.00   Min.   : 41.00
## Death   : 7915 1st Qu.: 60.50   1st Qu.: 74.50   1st Qu.: 74.50
##                Median : 68.50   Median : 85.50   Median : 83.50
##                Mean   : 69.32   Mean   : 86.66   Mean   : 84.76
##                3rd Qu.: 77.00   3rd Qu.: 97.50   3rd Qu.: 94.00
##                Max.   :127.50   Max.   :175.00   Max.   :148.00
##
## d1_glucose      d1_resprate      d1_sysbp      d1_temp

```

```
## Min. : 53.0 Min. : 7.00 Min. : 65.5 Min. :33.49
## 1st Qu.:109.5 1st Qu.:17.00 1st Qu.:109.0 1st Qu.:36.55
## Median :128.5 Median :19.50 Median :121.0 Median :36.76
## Mean :143.5 Mean :20.86 Mean :122.6 Mean :36.78
## 3rd Qu.:161.0 3rd Qu.:23.00 3rd Qu.:135.0 3rd Qu.:37.05
## Max. :449.5 Max. :96.00 Max. :196.0 Max. :38.85
##
## h1_diasbp h1_hearttrate h1_mbp h1_resprate
## Min. : 29.50 Min. : 41.00 Min. : 40.50 Min. : 10.00
## 1st Qu.: 58.50 1st Qu.: 73.50 1st Qu.: 73.00 1st Qu.: 32.00
## Median : 68.00 Median : 86.00 Median : 84.00 Median : 37.00
## Mean : 69.05 Mean : 87.89 Mean : 85.43 Mean : 39.71
## 3rd Qu.: 78.50 3rd Qu.:100.00 3rd Qu.: 96.00 3rd Qu.: 45.00
## Max. :128.00 Max. :154.00 Max. :151.50 Max. :248.00
##
## h1_spo2 h1_sysbp h1_sysbp_noninvasive
## Min. : 0.00 Min. : 64.0 Min. : 48.50
## 1st Qu.: 95.50 1st Qu.:107.0 1st Qu.: 85.50
## Median : 97.50 Median :123.0 Median : 96.00
## Mean : 96.65 Mean :124.7 Mean : 97.99
## 3rd Qu.: 99.00 3rd Qu.:140.0 3rd Qu.:109.00
## Max. :100.00 Max. :208.5 Max. :168.50
##
```

## Balancing Class

```
N_min <- min(unname(table(data$hospital_death)))
data <- ovun.sample(hospital_death ~ ., data=data, N=2*N_min, seed=1234)$data
```

## Splitting Data

```
set.seed(1234)
ind <- sample(2, nrow(data), replace = T, prob = c(0.8, 0.2))
trainData <- data[ind == 1,]
testData <- data[ind == 2,]
```

## Train Data

```
head(trainData,5)
```

```
## age bmi elective_surgery ethnicity gender height
## 1 72 27.99383 0 Caucasian M 180.0
## 2 73 27.95332 1 Caucasian F 163.8
## 3 68 37.25027 0 African American F 172.7
## 4 80 26.83518 0 Caucasian F 152.0
## 6 49 24.13270 0 Caucasian M 175.3
## icu_admit_source icu_stay_type icu_type pre_icu_los_days weight
```

## 1	Accident & Emergency	admit	MICU	0.165972222	90.70		
## 2	Operating Room / Recovery	admit	Med-Surg ICU	0.215277778	75.00		
## 3	Accident & Emergency	admit	CCU-CTICU	0.000694444	111.10		
## 4	Accident & Emergency	admit	Med-Surg ICU	0.223611111	62.00		
## 6	Accident & Emergency	admit	CCU-CTICU	0.004166667	74.16		
##	apache_2_diagnosis	apache_post_operative	arf_apache	gcs_eyes_apache			
## 1	113	0	0	3			
## 2	208	1	0	1			
## 3	113	0	0	4			
## 4	112	0	0	4			
## 6	113	0	0	4			
##	gcs_motor_apache	gcs_unable_apache	gcs_verbal_apache	heart_rate_apache			
## 1	5	0	1	97			
## 2	5	0	1	85			
## 3	6	0	5	106			
## 4	6	0	5	96			
## 6	6	0	5	115			
##	intubated_apache	map_apache	resprate_apache	temp_apache	ventilated_apache		
## 1	0	48	14	36.8	1		
## 2	1	130	5	36.6	1		
## 3	0	118	4	36.9	0		
## 4	0	127	41	37.0	0		
## 6	0	150	35	37.1	0		
##	d1_mbp_noninvasive_min	apache_4a_hospital_death_prob	apache_4a_icu_death_prob				
## 1	50	0.14	0.07				
## 2	55	0.29	0.21				
## 3	82	0.04	0.02				
## 4	85	0.07	0.03				
## 6	71	0.06	0.04				
##	aids	cirrhosis	diabetes_mellitus	hepatic_failure	immunosuppression	leukemia	
## 1	0	0	1	0	0	0	
## 2	0	0	0	0	0	0	
## 3	0	0	1	0	0	0	
## 4	0	0	1	0	0	0	
## 6	0	0	0	0	0	0	
##	lymphoma	solid_tumor_with_metastasis	apache_3j_bodysystem	hospital_death			
## 1	0	0	Sepsis	Survived			
## 2	0	0	Trauma	Survived			
## 3	0	0	Sepsis	Survived			
## 4	0	0	Cardiovascular	Survived			
## 6	0	0	Sepsis	Survived			
##	d1_diasbp	d1_heartrate	d1_mbp	d1_glucose	d1_resprate	d1_sysbp	d1_temp
## 1	61.5	90.5	74.5	255.5	22.0	117.0	37.40
## 2	65.0	70.0	75.0	138.5	14.0	117.5	36.75
## 3	72.0	93.0	100.0	163.5	22.0	148.5	37.35
## 4	59.0	65.5	93.5	215.0	22.5	117.0	37.00
## 6	79.5	104.0	90.5	129.0	27.0	130.5	37.30
##	h1_diasbp	h1_heartrate	h1_mbp	h1_resprate	h1_spo2	h1_sysbp	
## 1	49.0	91.0	66.5	43	98.5	103.0	
## 2	83.0	84.0	95.0	46	100.0	151.0	
## 3	81.0	97.0	98.5	47	98.5	163.0	
## 4	71.0	90.0	102.5	44	99.0	146.5	
## 6	84.5	109.5	95.5	67	95.5	133.0	
##	h1_sysbp_noninvasive						

```
## 1          79.0
## 2          117.0
## 3          125.0
## 4          113.5
## 6          110.5
```

## Test Data

```
head(testData ,5)
```

```
##      age      bmi elective_surgery      ethnicity gender height
## 5    66 40.14757              0      Caucasian      M  192.0
## 14   45 26.92346              0 Native American      M  172.7
## 16   67 27.65465              1      Caucasian      M  182.9
## 26   86 19.54726              1      Caucasian      F  152.4
## 28   83 27.86858              1      Caucasian      M  170.8
##      icu_admit_source icu_stay_type      icu_type pre_icu_los_days weight
## 5                Floor      admit Med-Surg ICU      0.7194444  148.0
## 14                Floor      admit Med-Surg ICU      5.7590278  80.3
## 16 Operating Room / Recovery      admit Med-Surg ICU      0.3194444  80.3
## 26 Operating Room / Recovery      admit      SICU      0.2638889  45.4
## 28 Operating Room / Recovery      admit      CSICU      2.3291667  81.3
##      apache_2_diagnosis apache_post_operative arf_apache gcs_eyes_apache
## 5                302              0              0              4
## 14                105              0              1              3
## 16                302              1              0              4
## 26                308              1              0              4
## 28                302              1              0              4
##      gcs_motor_apache gcs_unable_apache gcs_verbal_apache heart_rate_apache
## 5                5              0              1              113
## 14                5              0              4              100
## 16                6              0              5              113
## 26                6              0              5              118
## 28                6              0              5              60
##      intubated_apache map_apache resprate_apache temp_apache ventilated_apache
## 5                1          132              30      36.1              1
## 14                0           55              27      36.0              1
## 16                0          110              16      36.7              0
## 26                0          117              37      36.4              0
## 28                0           52              26      36.3              0
##      d1_mbp_noninvasive_min apache_4a_hospital_death_prob
## 5                63              0.52
## 14                55              0.08
## 16                88              0.05
## 26                88              0.02
## 28                52              0.06
##      apache_4a_icu_death_prob aids cirrhosis diabetes_mellitus hepatic_failure
## 5                0.32      0              0              0              0
## 14                0.05      0              0              1              0
## 16                0.02      0              0              0              0
## 26                0.01      0              0              0              0
## 28                0.02      0              0              0              0
```

```
##      immunosuppression leukemia lymphoma solid_tumor_with_metastasis
## 5              0          1          0                      0
## 14             0          0          0                      0
## 16             0          0          0                      0
## 26             0          0          0                      0
## 28             0          0          0                      0
##      apache_3j_bodysystem hospital_death d1_diasbp d1_hearttrate d1_mbp d1_glucose
## 5      Cardiovascular      Survived      86.0          88.5      93.5      176.5
## 14      Respiratory      Survived      73.5          85.0      78.0      100.5
## 16      Cardiovascular      Survived      77.0          92.5      99.0      128.5
## 26 Musculoskeletal/Skin      Survived      70.0          89.5      96.0      136.0
## 28      Cardiovascular      Survived      61.5          61.5      72.0      106.0
##      d1_resprate d1_sysbp d1_temp h1_diasbp h1_hearttrate h1_mbp h1_resprate
## 5          19.0      120.5      36.60      90.5          104.0      98.0          37
## 14          20.5       99.0      36.85      60.0           96.0      64.0          37
## 16          16.0      137.5      37.00      84.5           87.5     109.5          32
## 26          21.5      131.5      36.65      81.0          103.0     104.0          48
## 28          20.5      115.5      36.65      53.0           60.0      69.0          42
##      h1_spo2 h1_sysbp h1_sysbp_noninvasive
## 5          94.5      120.0              98.0
## 14          94.0       79.5              69.5
## 16          94.0      147.0             118.0
## 26          95.0      145.0             113.0
## 28          98.0      119.0             82.5
```

## Model Evaluation

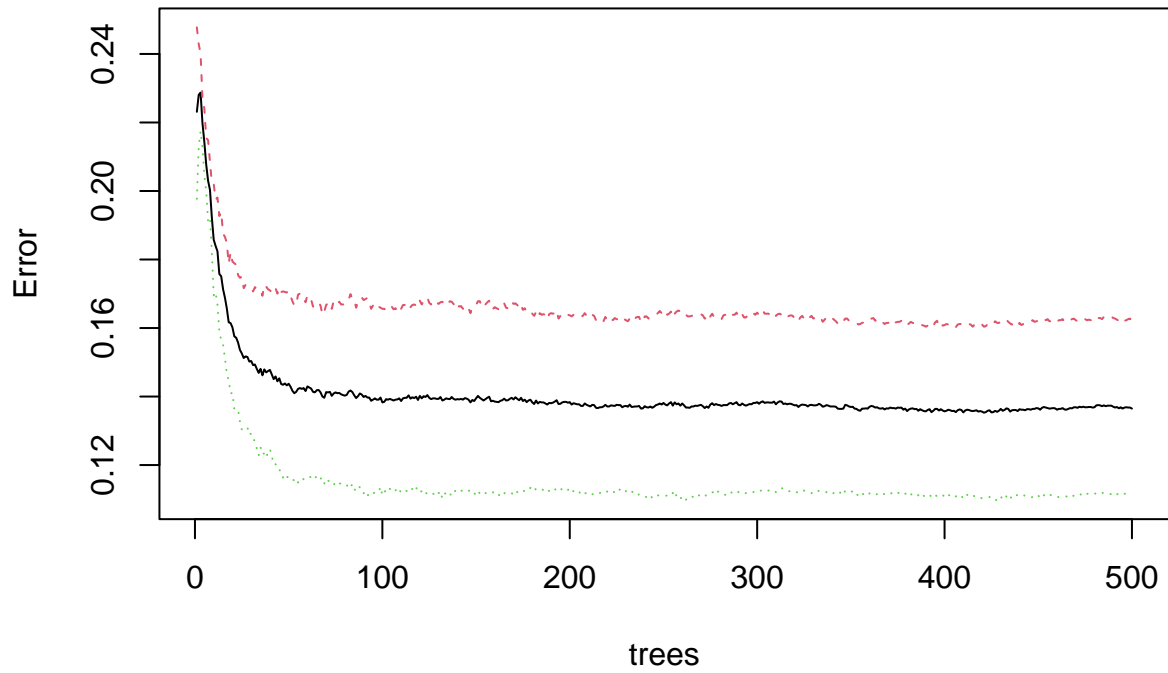
### Training and Plotting error

```
trainData_x <- trainData %>% select(-c(hospital_death))
trainData_y <- trainData$hospital_death

classifier_RF = randomForest(x = trainData_x,y = trainData_y)

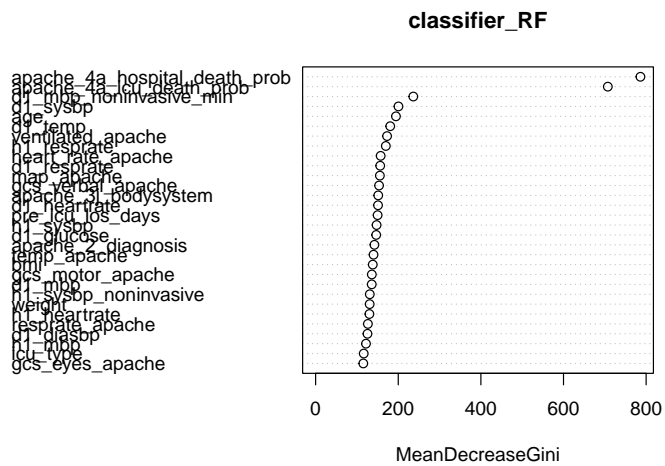
#classifier_RF = train(x = trainData_x,y = trainData_y)
plot(classifier_RF)
```

## classifier\_RF



## Plotting Variable Importance

```
varImpPlot(classifier_RF)
```



## Predict & Evaluation

```
testData_x <- testData %>% select(-c(hospital_death))
testData_y <- testData %>% select(c(hospital_death))
testData_y <- testData_y$hospital_death

predict_y <- predict(classifier_RF, testData_x)
predict_y <- as.vector(predict_y)
predict_y <- factor(predict_y, levels = c("Survived", "Death"))

# predict_y
confusionMatrix(predict_y, testData_y)
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction Survived Death
##   Survived      1310    176
##   Death         256   1426
##
##              Accuracy : 0.8636
##              95% CI : (0.8512, 0.8754)
##   No Information Rate : 0.5057
##   P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.7271
##
##   Mcnemar's Test P-Value : 0.0001442
##
##              Sensitivity : 0.8365
##              Specificity : 0.8901
##              Pos Pred Value : 0.8816
##              Neg Pred Value : 0.8478
##              Prevalence : 0.4943
##              Detection Rate : 0.4135
##              Detection Prevalence : 0.4691
##              Balanced Accuracy : 0.8633
##
##              'Positive' Class : Survived
##
```

## Tune Hyperparameter

```
ntreeku <- c()
mtryku <- c()
accuracyku <- c()
for(ntr in c(100,200,500,1000)){
  for(mt in c(7)){
    trainData_x <- trainData %>% select(-c(hospital_death))
    trainData_y <- trainData$hospital_death
```

```

classifier_RF = randomForest(x = trainData_x,y = trainData_y,
                             ntree=ntr,mtry = mt)
testData_x <- testData %>% select(-c(hospital_death))
testData_y <- testData %>% select(c(hospital_death))
testData_y <- testData_y$hospital_death

predict_y <- predict(classifier_RF,testData_x)
predict_y <- as.vector(predict_y)
predict_y <- factor(predict_y,levels = c("Survived","Death"))

# predict_y
acc <- confusionMatrix(predict_y,testData_y)$table
acc_c <- (acc[1,1] + acc[2,2])/sum(acc)
mtryku <- c(mtryku,mt)
accuracyku <- c(accuracyku,acc_c)
ntreeku <- c(ntreeku,ntr)
}
}

```

```

# mtry Number of variables randomly sampled as candidates at each split. Note that the default values are
table_hyper <- data.frame(ntreeku,mtryku,accuracyku)
table_hyper

```

```

##   ntreeku mtryku accuracyku
## 1     100      7  0.8636364
## 2     200      7  0.8655303
## 3     500      7  0.8661616
## 4    1000      7  0.8655303

```

```

trainData_x <- trainData %>% select(-c(hospital_death))
trainData_y <- trainData$hospital_death

classifier_RF = randomForest(x = trainData_x,y = trainData_y,
                             ntree=500,mtry = 7)
testData_x <- testData %>% select(-c(hospital_death))
testData_y <- testData %>% select(c(hospital_death))
testData_y <- testData_y$hospital_death

predict_y <- predict(classifier_RF,testData_x)
predict_y <- as.vector(predict_y)
predict_y <- factor(predict_y,levels = c("Survived","Death"))

# predict_y
confusionMatrix(predict_y,testData_y)

```

```

## Confusion Matrix and Statistics
##
##              Reference
## Prediction Survived Death
##   Survived    1314   176
##   Death       252  1426
##

```



```
##              Accuracy : 0.8649
##              95% CI : (0.8525, 0.8766)
##      No Information Rate : 0.5057
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.7296
##
##      McNemar's Test P-Value : 0.0002887
##
##              Sensitivity : 0.8391
##              Specificity : 0.8901
##      Pos Pred Value : 0.8819
##      Neg Pred Value : 0.8498
##              Prevalence : 0.4943
##      Detection Rate : 0.4148
##      Detection Prevalence : 0.4703
##      Balanced Accuracy : 0.8646
##
##      'Positive' Class : Survived
##
```

## Time Limit

```
current_2 <- Sys.time()
deltatime_rpart <- current_2 - current_0
print("Time Modelling :")
```

```
## [1] "Time Modelling :"
```

```
print(deltatime_rpart)
```

```
## Time difference of 4.959439 mins
```