

Covid19_StatisticalAnalysis

Safaa Dorian

2022-06-25

```
# Import the library
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(caret)
```

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

```
## Loading required package: lattice
```

```
# Read the data
df <- read.csv('clean_ml_data.csv', stringsAsFactors = F)
head(df)
```

```
##   gender pneumonia pregnant diabetes copd asthma immunosup hypertension
## 1      0          0          0          0      0      0          0          0
## 2      0          0          0          0      0      0          0          0
## 3      1          0          0          0      0      0          0          1
## 4      0          0          0          0      0      0          0          0
## 5      0          0          0          0      0      0          0          0
## 6      0          0          0          0      0      0          0          0
##   cardiovascular obesity renal_chronic tobacco death intubation ICU new_age
## 1              0      0              0      0      0              0  0      3
## 2              0      0              0      0      0              0  0      2
## 3              0      0              0      0      0              0  0      4
## 4              0      0              0      0      0              0  0      3
## 5              0      1              0      0      0              0  0      3
## 6              0      0              0      0      0              0  0      2
```

```
summary(df)
```

```
##      gender      pneumonia      pregnant      diabetes
## Min.   :0.0000   Min.   :0.0000   Min.   :0.000000   Min.   :0.0000
## 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.0000
## Median :0.0000   Median :0.0000   Median :0.000000   Median :0.0000
## Mean   :0.4946   Mean   :0.1587   Mean   :0.006572   Mean   :0.1413
## 3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.000000   3rd Qu.:0.0000
## Max.   :1.0000   Max.   :1.0000   Max.   :1.000000   Max.   :1.0000
##      copd      asthma      immunosup      hypertension
## Min.   :0.0000   Min.   :0.000000   Min.   :0.000000   Min.   :0.0000
## 1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.0000
## Median :0.0000   Median :0.000000   Median :0.000000   Median :0.0000
## Mean   :0.0123   Mean   :0.02402   Mean   :0.009475   Mean   :0.1813
## 3rd Qu.:0.0000   3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:0.0000
## Max.   :1.0000   Max.   :1.000000   Max.   :1.000000   Max.   :1.0000
## cardiovascular      obesity      renal_chronic      tobacco
## Min.   :0.000000   Min.   :0.0000   Min.   :0.000000   Min.   :0.000000
## 1st Qu.:0.000000   1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.000000
## Median :0.000000   Median :0.0000   Median :0.000000   Median :0.000000
## Mean   :0.01724   Mean   :0.1588   Mean   :0.01672   Mean   :0.07501
## 3rd Qu.:0.000000   3rd Qu.:0.0000   3rd Qu.:0.000000   3rd Qu.:0.000000
## Max.   :1.000000   Max.   :1.0000   Max.   :1.000000   Max.   :1.000000
##      death      intubation      ICU      new_age
## Min.   :0.0000   Min.   :0.000000   Min.   :0.000000   Min.   :1.00
## 1st Qu.:0.0000   1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:2.00
## Median :0.0000   Median :0.000000   Median :0.000000   Median :3.00
## Mean   :0.1032   Mean   :0.03247   Mean   :0.01771   Mean   :3.19
## 3rd Qu.:0.0000   3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:4.00
## Max.   :1.0000   Max.   :1.000000   Max.   :1.000000   Max.   :5.00
```

```
# transform type of columns from numerical to factors
```

```
df$gender = as.factor(df$gender)
df$new_age = as.factor(df$new_age)
df$pneumonia = as.factor(df$pneumonia)
df$pregnant = as.factor(df$pregnant)
df$diabetes = as.factor(df$diabetes)
df$copd = as.factor(df$copd)
df$asthma = as.factor(df$asthma)
df$immunosup = as.factor(df$immunosup)
df$hypertension = as.factor(df$hypertension)
df$cardiovascular = as.factor(df$cardiovascular)
df$obesity = as.factor(df$obesity)
df$renal_chronic = as.factor(df$renal_chronic)
df$tobacco = as.factor(df$tobacco)
df$death = as.factor(df$death)
df$intubation = as.factor(df$intubation)
df$ICU = as.factor(df$ICU)
```

```
summary(df)
```

```
## gender      pneumonia      pregnant      diabetes      copd      asthma
## 0:716505   0:1192767   0:1408374   0:1217403   0:1400255   0:1383639
## 1:701186   1: 224924   1: 9317    1: 200288   1: 17436    1: 34052
##
##
##
## immunosup      hypertension      cardiovascular      obesity      renal_chronic      tobacco
## 0:1404258   0:1160635   0:1393243   0:1192568   0:1393994   0:1311352
## 1: 13433   1: 257056   1: 24448   1: 225123   1: 23697   1: 106339
##
##
##
## death      intubation      ICU      new_age
## 0:1271359   0:1371652   0:1392579   1: 40886
## 1: 146332   1: 46039   1: 25112   2:361389
##                                     3:480422
##                                     4:357949
##                                     5:177045
```

Chi-squared test

```
# pregnancy risk
pregnancy_death <- table(df$pregnant,df$death) #generate contingency table with death column
chisq.test(pregnancy_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pregnancy_death  
## X-squared = 847.64, df = 1, p-value < 2.2e-16
```

```
pregnancy_ICU <- table(df$pregnant,df$ICU) #generate contingency table with ICU column  
chisq.test(pregnancy_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pregnancy_ICU  
## X-squared = 14.585, df = 1, p-value = 0.000134
```

```
pregnancy_intubation <- table(df$pregnant,df$intubation) #generate contingency table with intubation column  
chisq.test(pregnancy_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pregnancy_intubation  
## X-squared = 99.453, df = 1, p-value < 2.2e-16
```

```
# gender risk  
gender_death <- table(df$gender,df$death) #generate contingency table with death column  
chisq.test(gender_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: gender_death  
## X-squared = 10879, df = 1, p-value < 2.2e-16
```

```
gender_ICU <- table(df$gender,df$ICU) #generate contingency table with ICU column  
chisq.test(gender_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: gender_ICU  
## X-squared = 1809, df = 1, p-value < 2.2e-16
```

```
gender_intubation <- table(df$gender,df$intubation) #generate contingency table with intubation column  
chisq.test(gender_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: gender_intubation  
## X-squared = 3713.5, df = 1, p-value < 2.2e-16
```

```
# pneumonia risk  
pneumonia_death <- table(df$pneumonia,df$death) #generate contingency table with death column  
chisq.test(pneumonia_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pneumonia_death  
## X-squared = 393699, df = 1, p-value < 2.2e-16
```

```
pneumonia_ICU <- table(df$pneumonia,df$ICU) #generate contingency table with ICU column  
chisq.test(pneumonia_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pneumonia_ICU  
## X-squared = 97303, df = 1, p-value < 2.2e-16
```

```
pneumonia_intubation <- table(df$pneumonia,df$intubation) #generate contingency table with intubation column  
chisq.test(pneumonia_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: pneumonia_intubation  
## X-squared = 158699, df = 1, p-value < 2.2e-16
```

```
# diabetes risk  
diabetes_death <- table(df$diabetes,df$death) #generate contingency table with death column  
chisq.test(diabetes_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: diabetes_death  
## X-squared = 78160, df = 1, p-value < 2.2e-16
```

```
diabetes_ICU <- table(df$diabetes,df$ICU) #generate contingency table with ICU column  
chisq.test(diabetes_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: diabetes_ICU  
## X-squared = 7996.1, df = 1, p-value < 2.2e-16
```

```
diabetes_intubation <- table(df$diabetes,df$intubation) #generate contingency table with intubation column  
chisq.test(diabetes_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: diabetes_intubation  
## X-squared = 18387, df = 1, p-value < 2.2e-16
```

```
# copd risk  
copd_death <- table(df$copd,df$death) #generate contingency table with death column  
chisq.test(copd_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: copd_death  
## X-squared = 15085, df = 1, p-value < 2.2e-16
```

```
copd_ICU <- table(df$copd,df$ICU) #generate contingency table with ICU column  
chisq.test(copd_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: copd_ICU  
## X-squared = 1019.3, df = 1, p-value < 2.2e-16
```

```
copd_intubation <- table(df$copd,df$intubation) #generate contingency table with intubation column  
chisq.test(copd_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  copd_intubation  
## X-squared = 2898.1, df = 1, p-value < 2.2e-16
```

```
# asthma risk  
asthma_death <- table(df$asthma,df$death) #generate contingency table with death column  
chisq.test(asthma_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  asthma_death  
## X-squared = 195.9, df = 1, p-value < 2.2e-16
```

```
asthma_ICU <- table(df$asthma,df$ICU) #generate contingency table with ICU column  
chisq.test(asthma_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  asthma_ICU  
## X-squared = 6.3662, df = 1, p-value = 0.01163
```

```
asthma_intubation <- table(df$asthma,df$intubation) #generate contingency table with intubation column  
chisq.test(asthma_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  asthma_intubation  
## X-squared = 21.069, df = 1, p-value = 4.43e-06
```

```
# immunosup risk  
immunosup_death <- table(df$immunosup,df$death) #generate contingency table with death column  
chisq.test(immunosup_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  immunosup_death  
## X-squared = 3638.7, df = 1, p-value < 2.2e-16
```

```
immunosup_ICU <- table(df$immunosup,df$ICU) #generate contingency table with ICU column  
chisq.test(immunosup_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  immunosup_ICU  
## X-squared = 645.43, df = 1, p-value < 2.2e-16
```

```
immunosup_intubation <- table(df$immunosup,df$intubation) #generate contingency table with intubation column  
chisq.test(immunosup_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  immunosup_intubation  
## X-squared = 1254.7, df = 1, p-value < 2.2e-16
```

```
# hypertension risk  
hypertension_death <- table(df$hypertension,df$death) #generate contingency table with death column  
chisq.test(hypertension_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: hypertension_death  
## X-squared = 80997, df = 1, p-value < 2.2e-16
```

```
hypertension_ICU <- table(df$hypertension,df$ICU) #generate contingency table with ICU column  
chisq.test(hypertension_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: hypertension_ICU  
## X-squared = 6490.9, df = 1, p-value < 2.2e-16
```

```
hypertension_intubation <- table(df$hypertension,df$intubation) #generate contingency table with intubation column  
chisq.test(hypertension_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: hypertension_intubation  
## X-squared = 18418, df = 1, p-value < 2.2e-16
```

```
# cardiovascular risk  
cardiovascular_death <- table(df$cardiovascular,df$death) #generate contingency table with death column  
chisq.test(cardiovascular_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: cardiovascular_death  
## X-squared = 11490, df = 1, p-value < 2.2e-16
```

```
cardiovascular_ICU <- table(df$cardiovascular,df$ICU) #generate contingency table with ICU column  
chisq.test(cardiovascular_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: cardiovascular_ICU  
## X-squared = 1386.9, df = 1, p-value < 2.2e-16
```

```
cardiovascular_intubation <- table(df$cardiovascular,df$intubation) #generate contingency table with intubation column  
chisq.test(cardiovascular_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: cardiovascular_intubation  
## X-squared = 2605.8, df = 1, p-value < 2.2e-16
```

```
# obesity risk  
obesity_death <- table(df$obesity,df$death) #generate contingency table with death column  
chisq.test(obesity_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: obesity_death  
## X-squared = 6050.5, df = 1, p-value < 2.2e-16
```

```
obesity_ICU <- table(df$obesity,df$ICU) #generate contingency table with ICU column  
chisq.test(obesity_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: obesity_ICU  
## X-squared = 2562.6, df = 1, p-value < 2.2e-16
```

```
obesity_intubation <- table(df$obesity,df$intubation) #generate contingency table with intubation column  
chisq.test(obesity_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: obesity_intubation  
## X-squared = 3323.2, df = 1, p-value < 2.2e-16
```

```
# renal_chronic risk  
renal_chronic_death <- table(df$renal_chronic,df$death) #generate contingency table with death column  
chisq.test(renal_chronic_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: renal_chronic_death  
## X-squared = 29684, df = 1, p-value < 2.2e-16
```

```
renal_chronic_ICU <- table(df$renal_chronic,df$ICU) #generate contingency table with ICU column  
chisq.test(renal_chronic_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: renal_chronic_ICU  
## X-squared = 834.75, df = 1, p-value < 2.2e-16
```

```
renal_chronic_intubation <- table(df$renal_chronic,df$intubation) #generate contingency table with intubation column  
chisq.test(renal_chronic_intubation)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: renal_chronic_intubation  
## X-squared = 4584, df = 1, p-value < 2.2e-16
```

```
# tobacco risk  
tobacco_death <- table(df$tobacco,df$death) #generate contingency table with death column  
chisq.test(tobacco_death) #compare categorical distributions
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: tobacco_death  
## X-squared = 31.008, df = 1, p-value = 2.569e-08
```

```
tobacco_ICU <- table(df$tobacco,df$ICU) #generate contingency table with ICU column  
chisq.test(tobacco_ICU)
```

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: tobacco_ICU  
## X-squared = 0.027692, df = 1, p-value = 0.8678
```

```
tobacco_intubation <- table(df$tobacco,df$intubation) #generate contingency table with intubation column  
chisq.test(tobacco_intubation)
```

```
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data: tobacco_intubation
## X-squared = 62.694, df = 1, p-value = 2.415e-15
```

Logistic Regression to check the features significance

```
#make the model reproducible
set.seed(1)

#Use 70% of dataset as training set and remaining 30% as testing set
sample <- sample(c(TRUE, FALSE), nrow(df), replace=TRUE, prob=c(0.7,0.3))
train <- df[sample, ]
test <- df[!sample, ]
```

```
#fit logistic regression model with death column as outcome
model <- glm(death~ gender + pneumonia + pregnant + diabetes + copd + asthma + immunosup + hypertension + cardiovascular+
             obesity + renal_chronic + tobacco + new_age , family="binomial", data=train)
summary(model)
```

```
##
## Call:
## glm(formula = death ~ gender + pneumonia + pregnant + diabetes +
##      copd + asthma + immunosup + hypertension + cardiovascular +
##      obesity + renal_chronic + tobacco + new_age, family = "binomial",
##      data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5649  -0.3104  -0.1770  -0.1077   3.4190
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.765327   0.058869 -80.947 < 2e-16 ***
## gender1      -0.506296   0.008745 -57.895 < 2e-16 ***
## pneumonia1    2.598719   0.008520 305.028 < 2e-16 ***
## pregnant1    -0.175320   0.121484  -1.443  0.1490
## diabetes1     0.368457   0.009806  37.574 < 2e-16 ***
## copd1         0.300432   0.024495  12.265 < 2e-16 ***
## asthma1      -0.189303   0.030530  -6.201 5.63e-10 ***
## immunosup1    0.385857   0.031541  12.233 < 2e-16 ***
## hypertension1 0.253043   0.009697  26.095 < 2e-16 ***
## cardiovascular1 0.038416   0.022492   1.708  0.0876 .
## obesity1      0.257908   0.010470  24.633 < 2e-16 ***
## renal_chronic1 1.042593   0.021484  48.529 < 2e-16 ***
## tobacco1     -0.136388   0.016242  -8.397 < 2e-16 ***
## new_age2     -0.244596   0.062693  -3.901 9.56e-05 ***
## new_age3      0.864726   0.059546  14.522 < 2e-16 ***
## new_age4      1.894347   0.059225  31.985 < 2e-16 ***
## new_age5      2.907682   0.059373  48.973 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 658944  on 993231  degrees of freedom
## Residual deviance: 391539  on 993215  degrees of freedom
## AIC: 391573
##
## Number of Fisher Scoring iterations: 7
```

```
# feature importance
x = caret::varImp(model)
round(x,3)
```



```
## Overall
## gender1 57.895
## pneumonia1 305.028
## pregnant1 1.443
## diabetes1 37.574
## copd1 12.265
## asthma1 6.201
## immunosup1 12.233
## hypertension1 26.095
## cardiovascular1 1.708
## obesity1 24.633
## renal_chronic1 48.529
## tobacco1 8.397
## new_age2 3.901
## new_age3 14.522
## new_age4 31.985
## new_age5 48.973
```

```
#fit logistic regression model with icu column as outcome
model2 <- glm(ICU~ gender + pneumonia + pregnant + diabetes + copd + asthma + immunosup + hypertension + cardiovasc
ular+
              obesity + renal_chronic + tobacco + new_age , family="binomial", data=train)
summary(model2)
```

```
##
## Call:
## glm(formula = ICU ~ gender + pneumonia + pregnant + diabetes +
##      copd + asthma + immunosup + hypertension + cardiovascular +
##      obesity + renal_chronic + tobacco + new_age, family = "binomial",
##      data = train)
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1744  -0.0872  -0.0727  -0.0596   3.7080
##
```

```
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.819902  0.053015 -90.915 < 2e-16 ***
## gender1      -0.274692  0.016780 -16.370 < 2e-16 ***
## pneumonia1   3.457320  0.024394 141.729 < 2e-16 ***
## pregnant1    1.630286  0.091519  17.814 < 2e-16 ***
## diabetes1     0.127567  0.018412   6.929 4.25e-12 ***
## copd1        -0.031492  0.044732  -0.704 0.481420
## asthma1      -0.010256  0.055199  -0.186 0.852606
## immunosup1    0.281830  0.051962   5.424 5.83e-08 ***
## hypertension1 0.007575  0.018596   0.407 0.683755
## cardiovascular1 0.146379  0.038421   3.810 0.000139 ***
## obesity1     0.363925  0.018374  19.807 < 2e-16 ***
## renal_chronic1 -0.181887  0.041071  -4.429 9.49e-06 ***
## tobacco1     -0.129796  0.030292  -4.285 1.83e-05 ***
## new_age2     -1.649221  0.062075 -26.568 < 2e-16 ***
## new_age3     -1.115634  0.054690 -20.399 < 2e-16 ***
## new_age4     -0.840205  0.054162 -15.513 < 2e-16 ***
## new_age5     -0.751864  0.054920 -13.690 < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##      Null deviance: 177235  on 993231  degrees of freedom
## Residual deviance: 129891  on 993215  degrees of freedom
## AIC: 129925
##
## Number of Fisher Scoring iterations: 8
```

```
x2 = caret::varImp(model2)
round(x2,3)
```

```
## Overall
## gender1 16.370
## pneumonia1 141.729
## pregnant1 17.814
## diabetes1 6.929
## copd1 0.704
## asthma1 0.186
## immunosup1 5.424
## hypertension1 0.407
## cardiovascular1 3.810
## obesity1 19.807
## renal_chronic1 4.429
## tobacco1 4.285
## new_age2 26.568
## new_age3 20.399
## new_age4 15.513
## new_age5 13.690
```

#fit logistic regression model with intubation column as outcome

```
model3 <- glm(intubation~ gender + pneumonia + pregnant + diabetes + copd + asthma + immunosup + hypertension +
cardiovascular+
obesity + renal_chronic + tobacco + new_age , family="binomial", data=train)
summary(model3)
```

```
##
## Call:
## glm(formula = intubation ~ gender + pneumonia + pregnant + diabetes +
## copd + asthma + immunosup + hypertension + cardiovascular +
## obesity + renal_chronic + tobacco + new_age, family = "binomial",
## data = train)
##
```

Deviance Residuals:

```
## Min 1Q Median 3Q Max
## -1.0880 -0.1504 -0.1113 -0.0721 3.5601
##
```

Coefficients:

```
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.916847 0.059364 -82.825 < 2e-16 ***
## gender1 -0.338492 0.012853 -26.335 < 2e-16 ***
## pneumonia1 2.899283 0.015935 181.946 < 2e-16 ***
## pregnant1 0.790284 0.112002 7.056 1.71e-12 ***
## diabetes1 0.146395 0.013996 10.460 < 2e-16 ***
## copd1 0.052298 0.032707 1.599 0.1098
## asthma1 -0.039898 0.043429 -0.919 0.3583
## immunosup1 0.302159 0.041007 7.369 1.73e-13 ***
## hypertension1 0.134339 0.014006 9.592 < 2e-16 ***
## cardiovascular1 0.033028 0.029892 1.105 0.2692
## obesity1 0.285308 0.014468 19.720 < 2e-16 ***
## renal_chronic1 0.170784 0.028387 6.016 1.78e-09 ***
## tobacco1 0.006578 0.022267 0.295 0.7677
## new_age2 -1.040161 0.066815 -15.568 < 2e-16 ***
## new_age3 -0.164904 0.060707 -2.716 0.0066 **
## new_age4 0.440718 0.060158 7.326 2.37e-13 ***
## new_age5 0.672691 0.060556 11.109 < 2e-16 ***
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
## Null deviance: 284885 on 993231 degrees of freedom
```

```
## Residual deviance: 203566 on 993215 degrees of freedom
```

```
## AIC: 203600
```

```
##
```

```
## Number of Fisher Scoring iterations: 8
```

```
x3 = caret::varImp(model3)
round(x3,3)
```

##	Overall
## gender1	26.335
## pneumonia1	181.946
## pregnant1	7.056
## diabetes1	10.460
## copd1	1.599
## asthma1	0.919
## immunosup1	7.369
## hypertension1	9.592
## cardiovascular1	1.105
## obesity1	19.720
## renal_chronic1	6.016
## tobacco1	0.295
## new_age2	15.568
## new_age3	2.716
## new_age4	7.326
## new_age5	11.109

...