

# 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.   :0.00000 Min.   :0.000000 Min.   :0.0000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.000000 1st Qu.:0.0000
## Median :0.0000 Median :0.00000 Median :0.000000 Median :0.0000
## Mean   :0.0123 Mean   :0.02402 Mean   :0.009475 Mean   :0.1813
## 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.000000 3rd Qu.:0.0000
## Max.   :1.0000 Max.   :1.00000 Max.   :1.000000 Max.   :1.0000
## cardiovascular obesity      renal_chronic      tobacco
## Min.   :0.00000 Min.   :0.0000 Min.   :0.00000 Min.   :0.00000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000
## Median :0.00000 Median :0.0000 Median :0.00000 Median :0.00000
## Mean   :0.01724 Mean   :0.1588 Mean   :0.01672 Mean   :0.07501
## 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000
## Max.   :1.00000 Max.   :1.0000 Max.   :1.00000 Max.   :1.00000
##      death      intubation      ICU      new_age
## Min.   :0.0000 Min.   :0.00000 Min.   :0.00000 Min.   :1.00
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:2.00
## Median :0.0000 Median :0.00000 Median :0.00000 Median :3.00
## Mean   :0.1032 Mean   :0.03247 Mean   :0.01771 Mean   :3.19
## 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:4.00
## Max.   :1.0000 Max.   :1.00000 Max.   :1.00000 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$pregnant = as.factor(df$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 Min.   :0.0000 0:1408374 0:1217403 0:1400255 0:1383639
## 1:701186 1st Qu.:0.0000 1: 9317 1: 200288 1: 17436 1: 34052
##      Median :0.0000
##      Mean   :0.1587
##      3rd Qu.:0.0000
##      Max.   :1.0000
## 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
```

```
# 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+ pregnant + diabetes + copd + asthma + immunosup + hypertension + cardiovascular+  
            obesity + renal_chronic + tobacco + new_age , family="binomial", data=train)  
summary(model)
```

```
##
## Call:
## glm(formula = death ~ gender + pregnant + diabetes + copd + asthma +
##      immunosup + hypertension + cardiovascular + obesity + renal_chronic +
##      tobacco + new_age, family = "binomial", data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3876  -0.3961  -0.2107  -0.1387   3.3210
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.312361   0.057787 -74.626 < 2e-16 ***
## gender1      -0.648665   0.007703 -84.207 < 2e-16 ***
## pregnant1    -0.010161   0.119333  -0.085  0.932
## diabetes1     0.551486   0.008592  64.187 < 2e-16 ***
## copd1         0.436656   0.021267  20.532 < 2e-16 ***
## asthma1      -0.222535   0.027418  -8.116 4.80e-16 ***
## immunosup1    0.571703   0.028290  20.209 < 2e-16 ***
## hypertension1 0.314929   0.008481  37.133 < 2e-16 ***
## cardiovascular1 0.106797   0.019565   5.459 4.80e-08 ***
## obesity1      0.398728   0.009252  43.098 < 2e-16 ***
## renal_chronic1 1.160111   0.019218  60.367 < 2e-16 ***
## tobacco1     -0.086094   0.014258  -6.038 1.56e-09 ***
## new_age2     -0.240754   0.061649  -3.905 9.41e-05 ***
## new_age3      1.163888   0.058465  19.907 < 2e-16 ***
## new_age4      2.455231   0.058126  42.240 < 2e-16 ***
## new_age5      3.648835   0.058222  62.672 < 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: 494146  on 993216  degrees of freedom
## AIC: 494178
##
## Number of Fisher Scoring iterations: 7
```

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

```
##              Overall
## gender1          84.207
## pregnant1         0.085
## diabetes1        64.187
## copd1            20.532
## asthma1           8.116
## immunosup1       20.209
## hypertension1    37.133
## cardiovascular1  5.459
## obesity1         43.098
## renal_chronic1   60.367
## tobacco1         6.038
## new_age2          3.905
## new_age3         19.907
## new_age4         42.240
## new_age5         62.672
```

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



```
##
## Call:
## glm(formula = ICU ~ gender + pregnant + diabetes + copd + asthma +
##      immunosup + hypertension + cardiovascular + obesity + renal_chronic +
##      tobacco + new_age, family = "binomial", data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8330  -0.2209  -0.1463  -0.0971   3.5361
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.99827    0.04892 -81.729 < 2e-16 ***
## gender1       -0.55995    0.01619 -34.578 < 2e-16 ***
## pregnant1      1.86796    0.08530  21.898 < 2e-16 ***
## diabetes1      0.43995    0.01826  24.097 < 2e-16 ***
## copd1          0.15646    0.04409   3.549 0.000387 ***
## asthma1       -0.07660    0.05371  -1.426 0.153801
## immunosup1     0.58942    0.05079  11.604 < 2e-16 ***
## hypertension1  0.14470    0.01840   7.864 3.72e-15 ***
## cardiovascular1 0.21231    0.03765   5.638 1.72e-08 ***
## obesity1       0.58407    0.01784  32.746 < 2e-16 ***
## renal_chronic1 0.13571    0.04078   3.328 0.000875 ***
## tobacco1      -0.12374    0.02951  -4.194 2.75e-05 ***
## new_age2       -1.56796    0.05947 -26.365 < 2e-16 ***
## new_age3       -0.53366    0.05167 -10.327 < 2e-16 ***
## new_age4        0.29709    0.05079   5.849 4.95e-09 ***
## new_age5        0.85561    0.05156  16.594 < 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: 162048  on 993216  degrees of freedom
## AIC: 162080
##
## Number of Fisher Scoring iterations: 8
```

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

```
##              Overall
## gender1         34.578
## pregnant1       21.898
## diabetes1       24.097
## copd1            3.549
## asthma1         1.426
## immunosup1     11.604
## hypertension1   7.864
## cardiovascular1 5.638
## obesity1       32.746
## renal_chronic1  3.328
## tobacco1        4.194
## new_age2       26.365
## new_age3       10.327
## new_age4        5.849
## new_age5       16.594
```

```
#fit logistic regression model with intubation column as outcome
model3 <- glm(intubation~ gender + pregnant + diabetes + copd + asthma + immunosup + hypertension + cardiovascular +
r+
              obesity + renal_chronic + tobacco + new_age , family="binomial", data=train)
summary(model3)
```

```
##
## Call:
## glm(formula = intubation ~ gender + pregnant + diabetes + copd +
##      asthma + immunosup + hypertension + cardiovascular + obesity +
##      renal_chronic + tobacco + new_age, family = "binomial", data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2086  -0.3073  -0.1827  -0.0972   3.4755
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.32885    0.05769  -75.038 < 2e-16 ***
## gender1       -0.57723    0.01220  -47.309 < 2e-16 ***
## pregnant1      1.01871    0.10869   9.373 < 2e-16 ***
## diabetes1      0.41348    0.01348  30.663 < 2e-16 ***
## copd1          0.22740    0.03148   7.224 5.05e-13 ***
## asthma1       -0.10685    0.04168  -2.563 0.01037 *
## immunosup1     0.53078    0.03944  13.459 < 2e-16 ***
## hypertension1  0.24853    0.01349  18.425 < 2e-16 ***
## cardiovascular1 0.09366    0.02867   3.267 0.00109 **
## obesity1       0.46540    0.01378  33.770 < 2e-16 ***
## renal_chronic1 0.44923    0.02743  16.376 < 2e-16 ***
## tobacco1       0.01710    0.02121   0.806 0.42014
## new_age2       -1.02429    0.06555 -15.626 < 2e-16 ***
## new_age3        0.24382    0.05927   4.114 3.90e-05 ***
## new_age4        1.29954    0.05856  22.193 < 2e-16 ***
## new_age5        1.95627    0.05887  33.231 < 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: 248754  on 993216  degrees of freedom
## AIC: 248786
##
## Number of Fisher Scoring iterations: 8
```

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

```
##              Overall
## gender1         47.309
## pregnant1        9.373
## diabetes1       30.663
## copd1            7.224
## asthma1          2.563
## immunosup1      13.459
## hypertension1   18.425
## cardiovascular1  3.267
## obesity1        33.770
## renal_chronic1  16.376
## tobacco1         0.806
## new_age2        15.626
## new_age3         4.114
## new_age4        22.193
## new_age5        33.231
```

...