> setwd('C:/Rdirectory/data\_mining/data\_mining\_covid');

> covid\_original = read.csv("covid\_original.csv", header=T);

> covid\_original$sex = as.factor(covid\_original$sex);

> covid\_original$patient\_type = as.factor(covid\_original$patient\_type);

> covid\_original$intubated = as.factor(covid\_original$intubated);

> covid\_original$pneumonia = as.factor(covid\_original$pneumonia);

> covid\_original$age = as.numeric(covid\_original$age);

> covid\_original$pregnancy = as.factor(covid\_original$pregnancy);

> covid\_original$diabetes = as.factor(covid\_original$diabetes);

> covid\_original$copd = as.factor(covid\_original$copd);

> covid\_original$asthma = as.factor(covid\_original$asthma);

> covid\_original$inmsupr = as.factor(covid\_original$inmsupr);

> covid\_original$hypertension = as.factor(covid\_original$hypertension);

> covid\_original$other\_disease = as.factor(covid\_original$other\_disease);

> covid\_original$cardiovascular = as.factor(covid\_original$cardiovascular);

> covid\_original$obesity = as.factor(covid\_original$obesity);

> covid\_original$renal\_chronic = as.factor(covid\_original$renal\_chronic);

> covid\_original$tobacco = as.factor(covid\_original$tobacco);

> covid\_original$contact\_other\_covid = as.factor(covid\_original$contact\_other\_covid);

> covid\_original$covid\_res = as.factor(covid\_original$covid\_res);

> covid\_original$icu = as.factor(covid\_original$icu);

> covid\_original$is\_dead = as.factor(covid\_original$is\_dead);

> covid\_original$day\_cnt = as.numeric(covid\_original$day\_cnt);

> str(covid\_original);

'data.frame': 566602 obs. of 25 variables:

$ id : chr "16169f" "1009bf" "167386" "0b5948" ...

$ sex : Factor w/ 2 levels "1","2": 2 2 1 2 1 2 2 1 1 1 ...

$ patient\_type : Factor w/ 2 levels "1","2": 1 1 2 2 2 2 2 1 1 2 ...

$ entry\_date : chr "04-05-2020" "19-03-2020" "06-04-2020" "17-04-2020" ...

$ date\_symptoms : chr "02-05-2020" "17-03-2020" "01-04-2020" "10-04-2020" ...

$ date\_died : chr "9999-99-99" "9999-99-99" "9999-99-99" "9999-99-99" ...

$ intubated : Factor w/ 4 levels "1","2","97","99": 3 3 2 2 2 2 2 3 3 1 ...

$ pneumonia : Factor w/ 3 levels "1","2","99": 2 2 2 1 2 1 2 2 2 1 ...

$ age : num 27 24 54 30 60 47 63 56 41 39 ...

$ pregnancy : Factor w/ 4 levels "1","2","97","98": 3 3 2 3 2 3 3 2 2 2 ...

$ diabetes : Factor w/ 3 levels "1","2","98": 2 2 2 2 1 1 2 2 2 2 ...

$ copd : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 2 2 2 ...

$ asthma : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 2 2 2 ...

$ inmsupr : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 2 2 2 ...

$ hypertension : Factor w/ 3 levels "1","2","98": 2 2 2 2 1 2 1 1 2 2 ...

$ other\_disease : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 2 2 2 ...

$ cardiovascular : Factor w/ 3 levels "1","2","98": 2 2 2 2 1 2 2 2 2 2 ...

$ obesity : Factor w/ 3 levels "1","2","98": 2 2 1 2 2 2 2 2 2 1 ...

$ renal\_chronic : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 1 2 2 ...

$ tobacco : Factor w/ 3 levels "1","2","98": 2 2 2 2 2 2 2 1 2 2 ...

$ contact\_other\_covid: Factor w/ 3 levels "1","2","99": 2 3 3 3 3 3 3 1 3 3 ...

$ covid\_res : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...

$ icu : Factor w/ 4 levels "1","2","97","99": 3 3 2 2 2 1 2 3 3 2 ...

$ is\_dead : Factor w/ 2 levels "1","2": 2 2 2 2 1 1 2 2 2 2 ...

$ day\_cnt : num NA NA NA NA 9 13 NA NA NA NA ...

> covid\_original\_without\_missing = covid\_original[

+ (covid\_original$intubated == 1 | covid\_original$intubated == 2)

+ & (covid\_original$pneumonia == 1 | covid\_original$pneumonia == 2)

+ & (covid\_original$pregnancy == 1 | covid\_original$pregnancy == 2)

+ & (covid\_original$diabetes == 1 | covid\_original$diabetes == 2)

+ & (covid\_original$copd == 1 | covid\_original$copd == 2)

+ & (covid\_original$asthma == 1 | covid\_original$asthma == 2)

+ & (covid\_original$inmsupr == 1 | covid\_original$inmsupr == 2)

+ & (covid\_original$hypertension == 1 | covid\_original$hypertension == 2)

+ & (covid\_original$other\_disease == 1 | covid\_original$other\_disease == 2)

+ & (covid\_original$cardiovascular == 1 | covid\_original$cardiovascular == 2)

+ & (covid\_original$obesity == 1 | covid\_original$obesity == 2)

+ & (covid\_original$renal\_chronic == 1 | covid\_original$renal\_chronic == 2)

+ & (covid\_original$tobacco == 1 | covid\_original$tobacco == 2)

+ & (covid\_original$contact\_other\_covid == 1 | covid\_original$contact\_other\_covid == 2)

+ & (covid\_original$icu == 1 | covid\_original$icu == 2)

+ ,];

> nrow(covid\_original\_without\_missing); # 23158

[1] 23158

> covid\_original\_dead\_without\_missing = covid\_original\_without\_missing[covid\_original\_without\_missing$is\_dead==1,];

> nrow(covid\_original\_dead\_without\_missing); # 4020

[1] 4020

> covid\_original\_without\_missing = covid\_original\_without\_missing[,!names(covid\_original\_without\_missing) %in% c("id", "entry\_date","date\_symptoms", "date\_died", "day\_cnt", "intubated", "pregnancy", "covid\_res", "icu")];

> covid\_train = covid\_original\_without\_missing[1:20000,]; # 20000개

> covid\_test = covid\_original\_without\_missing[-(1:20000),]; #3158개

> # install.packages("ROSE");

> library(ROSE);

> install.packages("ROSE");

WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/

‘C:/Users/JNP/Documents/R/win-library/4.1’의 위치에 패키지(들)을 설치합니다.

(왜냐하면 ‘lib’가 지정되지 않았기 때문입니다)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/ROSE\_0.0-4.zip'

Content type 'application/zip' length 113607 bytes (110 KB)

downloaded 110 KB

패키지 ‘ROSE’를 성공적으로 압축해제하였고 MD5 sums 이 확인되었습니다

다운로드된 바이너리 패키지들은 다음의 위치에 있습니다

C:\Users\JNP\AppData\Local\Temp\RtmpmcFquu\downloaded\_packages

> #install.packages("ROSE");

> library(ROSE);

Loaded ROSE 0.0-4

> train\_oversampling\_cnt = nrow(covid\_train) - 2\*nrow(covid\_train[covid\_train$is\_dead == 1, ]); #12356

> train\_oversampling\_data = ROSE(is\_dead ~ ., data = covid\_train, N=50000, seed=10 )$data;

> train\_oversampling\_data\_dead = train\_oversampling\_data[train\_oversampling\_data$is\_dead==1,];

> covid\_train\_oversampled = rbind(covid\_train, train\_oversampling\_data\_dead[1:train\_oversampling\_cnt,]); #32356개

> write.csv(covid\_train\_oversampled, "covid\_train.csv", row.names = F);

> test\_oversampling\_cnt = nrow(covid\_test) - 2\*nrow(covid\_test[covid\_test$is\_dead == 1, ]); #2762

> test\_oversampling\_data = ROSE(is\_dead ~ ., data = covid\_test, N=50000, seed=20 )$data;

> test\_oversampling\_data\_dead = test\_oversampling\_data[test\_oversampling\_data$is\_dead==1,];

> covid\_test\_oversampled = rbind(covid\_test, test\_oversampling\_data\_dead[1:test\_oversampling\_cnt,]); #5920개

> write.csv(covid\_test\_oversampled, "covid\_test.csv", row.names = F);

> nrow(covid\_test\_oversampled[covid\_test\_oversampled$is\_dead==2,])

[1] 2960

> covid\_original\_dead\_without\_missing = covid\_original\_dead\_without\_missing[,!names(covid\_original\_dead\_without\_missing) %in% c("id", "entry\_date","date\_symptoms", "date\_died", "is\_dead", "intubated", "pregnancy", "covid\_res", "icu")];

> covid\_dead\_train = covid\_original\_dead\_without\_missing[1:3600,];

> covid\_dead\_test = covid\_original\_dead\_without\_missing[-(1:3600),];

> write.csv(covid\_dead\_train, "covid\_dead\_train.csv", row.names = F);

> write.csv(covid\_dead\_test, "covid\_dead\_test.csv", row.names = F);

> covid\_train = read.csv("covid\_train.csv", header=T);

> covid\_train$sex = as.factor(covid\_train$sex);

> covid\_train$patient\_type = as.factor(covid\_train$patient\_type);

> covid\_train$pneumonia = as.factor(covid\_train$pneumonia);

> covid\_train$age = as.numeric(covid\_train$age);

> covid\_train$diabetes = as.factor(covid\_train$diabetes);

> covid\_train$copd = as.factor(covid\_train$copd);

> covid\_train$asthma = as.factor(covid\_train$asthma);

> covid\_train$inmsupr = as.factor(covid\_train$inmsupr);

> covid\_train$hypertension = as.factor(covid\_train$hypertension);

> covid\_train$other\_disease = as.factor(covid\_train$other\_disease);

> covid\_train$cardiovascular = as.factor(covid\_train$cardiovascular);

> covid\_train$obesity = as.factor(covid\_train$obesity);

> covid\_train$renal\_chronic = as.factor(covid\_train$renal\_chronic);

> covid\_train$tobacco = as.factor(covid\_train$tobacco);

> covid\_train$contact\_other\_covid = as.factor(covid\_train$contact\_other\_covid);

> covid\_train$is\_dead = as.factor(covid\_train$is\_dead);

> str(covid\_train);

'data.frame': 32356 obs. of 16 variables:

$ sex : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

$ patient\_type : Factor w/ 1 level "2": 1 1 1 1 1 1 1 1 1 1 ...

$ pneumonia : Factor w/ 2 levels "1","2": 2 2 1 1 1 2 1 1 1 2 ...

$ age : num 25 52 51 67 59 52 54 78 80 40 ...

$ diabetes : Factor w/ 2 levels "1","2": 2 2 2 1 1 1 2 2 1 2 ...

$ copd : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ asthma : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ inmsupr : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ hypertension : Factor w/ 2 levels "1","2": 2 2 2 1 2 1 2 1 1 2 ...

$ other\_disease : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ cardiovascular : Factor w/ 2 levels "1","2": 2 2 2 2 2 1 2 2 2 2 ...

$ obesity : Factor w/ 2 levels "1","2": 2 1 2 1 2 2 2 1 1 2 ...

$ renal\_chronic : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ tobacco : Factor w/ 2 levels "1","2": 2 1 2 2 2 2 2 2 2 2 ...

$ contact\_other\_covid: Factor w/ 2 levels "1","2": 1 1 1 1 1 2 2 1 1 2 ...

$ is\_dead : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 1 2 2 ...

> head(covid\_train);

sex patient\_type pneumonia age diabetes copd asthma inmsupr hypertension other\_disease cardiovascular obesity renal\_chronic tobacco

1 1 2 2 25 2 2 2 2 2 2 2 2 2 2

2 1 2 2 52 2 2 2 2 2 2 2 1 2 1

3 1 2 1 51 2 2 2 2 2 2 2 2 2 2

4 1 2 1 67 1 2 2 2 1 2 2 1 2 2

5 1 2 1 59 1 2 2 2 2 2 2 2 2 2

6 1 2 2 52 1 2 2 2 1 2 1 2 2 2

contact\_other\_covid is\_dead

1 1 2

2 1 2

3 1 2

4 1 2

5 1 2

6 2 2

> covid\_test = read.csv("covid\_test.csv", header=T);

> covid\_test$sex = as.factor(covid\_test$sex);

> covid\_test$patient\_type = as.factor(covid\_test$patient\_type);

> covid\_test$pneumonia = as.factor(covid\_test$pneumonia);

> covid\_test$age = as.numeric(covid\_test$age);

> covid\_test$diabetes = as.factor(covid\_test$diabetes);

> covid\_test$copd = as.factor(covid\_test$copd);

> covid\_test$asthma = as.factor(covid\_test$asthma);

> covid\_test$inmsupr = as.factor(covid\_test$inmsupr);

> covid\_test$hypertension = as.factor(covid\_test$hypertension);

> covid\_test$other\_disease = as.factor(covid\_test$other\_disease);

> covid\_test$cardiovascular = as.factor(covid\_test$cardiovascular);

> covid\_test$obesity = as.factor(covid\_test$obesity);

> covid\_test$renal\_chronic = as.factor(covid\_test$renal\_chronic);

> covid\_test$tobacco = as.factor(covid\_test$tobacco);

> covid\_test$contact\_other\_covid = as.factor(covid\_test$contact\_other\_covid);

> covid\_test$is\_dead = as.factor(covid\_test$is\_dead);

> str(covid\_test);

'data.frame': 5920 obs. of 16 variables:

$ sex : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

$ patient\_type : Factor w/ 1 level "2": 1 1 1 1 1 1 1 1 1 1 ...

$ pneumonia : Factor w/ 2 levels "1","2": 2 2 1 1 1 1 2 2 2 1 ...

$ age : num 52 36 0 85 75 19 46 60 62 39 ...

$ diabetes : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 1 2 ...

$ copd : Factor w/ 2 levels "1","2": 2 2 2 1 2 2 2 2 2 2 ...

$ asthma : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ inmsupr : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ hypertension : Factor w/ 2 levels "1","2": 2 2 2 1 1 2 2 2 2 1 ...

$ other\_disease : Factor w/ 2 levels "1","2": 1 2 2 2 2 2 2 2 2 2 ...

$ cardiovascular : Factor w/ 2 levels "1","2": 2 2 2 1 2 2 2 2 2 2 ...

$ obesity : Factor w/ 2 levels "1","2": 2 2 2 1 2 2 2 2 2 2 ...

$ renal\_chronic : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ tobacco : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ contact\_other\_covid: Factor w/ 2 levels "1","2": 2 2 2 2 2 2 1 2 2 2 ...

$ is\_dead : Factor w/ 2 levels "1","2": 2 2 1 2 2 2 2 2 2 2 ...

> head(covid\_test);

sex patient\_type pneumonia age diabetes copd asthma inmsupr hypertension other\_disease cardiovascular obesity renal\_chronic tobacco

1 1 2 2 52 2 2 2 2 2 1 2 2 2 2

2 1 2 2 36 2 2 2 2 2 2 2 2 2 2

3 1 2 1 0 2 2 2 2 2 2 2 2 2 2

4 1 2 1 85 2 1 2 2 1 2 1 1 2 2

5 1 2 1 75 2 2 2 2 1 2 2 2 2 2

6 1 2 1 19 2 2 2 2 2 2 2 2 2 2

contact\_other\_covid is\_dead

1 2 2

2 2 2

3 2 1

4 2 2

5 2 2

6 2 2

> covid\_dead\_train = read.csv("covid\_dead\_train.csv", header=T);

> covid\_dead\_train$sex = as.factor(covid\_dead\_train$sex);

> covid\_dead\_train$patient\_type = as.factor(covid\_dead\_train$patient\_type);

> covid\_dead\_train$pneumonia = as.factor(covid\_dead\_train$pneumonia);

> covid\_dead\_train$age = as.numeric(covid\_dead\_train$age);

> covid\_dead\_train$diabetes = as.factor(covid\_dead\_train$diabetes);

> covid\_dead\_train$copd = as.factor(covid\_dead\_train$copd);

> covid\_dead\_train$asthma = as.factor(covid\_dead\_train$asthma);

> covid\_dead\_train$inmsupr = as.factor(covid\_dead\_train$inmsupr);

> covid\_dead\_train$hypertension = as.factor(covid\_dead\_train$hypertension);

> covid\_dead\_train$other\_disease = as.factor(covid\_dead\_train$other\_disease);

> covid\_dead\_train$cardiovascular = as.factor(covid\_dead\_train$cardiovascular);

> covid\_dead\_train$obesity = as.factor(covid\_dead\_train$obesity);

> covid\_dead\_train$renal\_chronic = as.factor(covid\_dead\_train$renal\_chronic);

> covid\_dead\_train$tobacco = as.factor(covid\_dead\_train$tobacco);

> covid\_dead\_train$contact\_other\_covid = as.factor(covid\_dead\_train$contact\_other\_covid);

> covid\_dead\_train$day\_cnt = as.numeric(covid\_dead\_train$day\_cnt);

> str(covid\_dead\_train);

'data.frame': 3600 obs. of 16 variables:

$ sex : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

$ patient\_type : Factor w/ 1 level "2": 1 1 1 1 1 1 1 1 1 1 ...

$ pneumonia : Factor w/ 2 levels "1","2": 1 1 2 1 1 1 1 1 1 2 ...

$ age : num 78 65 58 78 92 70 71 65 56 72 ...

$ diabetes : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 1 1 2 2 ...

$ copd : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 1 2 2 2 ...

$ asthma : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ inmsupr : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ hypertension : Factor w/ 2 levels "1","2": 1 2 2 2 1 2 1 1 2 1 ...

$ other\_disease : Factor w/ 2 levels "1","2": 2 2 1 2 2 2 2 2 2 2 ...

$ cardiovascular : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ obesity : Factor w/ 2 levels "1","2": 1 1 2 1 2 1 1 1 2 2 ...

$ renal\_chronic : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 1 ...

$ tobacco : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 1 2 2 2 ...

$ contact\_other\_covid: Factor w/ 2 levels "1","2": 1 2 1 1 2 2 1 2 2 2 ...

$ day\_cnt : num 10 7 1 6 3 16 21 14 30 20 ...

> head(covid\_dead\_train);

sex patient\_type pneumonia age diabetes copd asthma inmsupr hypertension other\_disease cardiovascular obesity renal\_chronic tobacco

1 1 2 1 78 2 2 2 2 1 2 2 1 2 2

2 1 2 1 65 2 2 2 2 2 2 2 1 2 2

3 1 2 2 58 2 2 2 2 2 1 2 2 2 2

4 1 2 1 78 2 2 2 2 2 2 2 1 2 2

5 1 2 1 92 2 2 2 2 1 2 2 2 2 2

6 1 2 1 70 2 2 2 2 2 2 2 1 2 2

contact\_other\_covid day\_cnt

1 1 10

2 2 7

3 1 1

4 1 6

5 2 3

6 2 16

> covid\_dead\_test = read.csv("covid\_dead\_test.csv", header=T);

> covid\_dead\_test$sex = as.factor(covid\_dead\_test$sex);

> covid\_dead\_test$patient\_type = as.factor(covid\_dead\_test$patient\_type);

> covid\_dead\_test$pneumonia = as.factor(covid\_dead\_test$pneumonia);

> covid\_dead\_test$age = as.numeric(covid\_dead\_test$age);

> covid\_dead\_test$diabetes = as.factor(covid\_dead\_test$diabetes);

> covid\_dead\_test$copd = as.factor(covid\_dead\_test$copd);

> covid\_dead\_test$asthma = as.factor(covid\_dead\_test$asthma);

> covid\_dead\_test$inmsupr = as.factor(covid\_dead\_test$inmsupr);

> covid\_dead\_test$hypertension = as.factor(covid\_dead\_test$hypertension);

> covid\_dead\_test$other\_disease = as.factor(covid\_dead\_test$other\_disease);

> covid\_dead\_test$cardiovascular = as.factor(covid\_dead\_test$cardiovascular);

> covid\_dead\_test$obesity = as.factor(covid\_dead\_test$obesity);

> covid\_dead\_test$renal\_chronic = as.factor(covid\_dead\_test$renal\_chronic);

> covid\_dead\_test$tobacco = as.factor(covid\_dead\_test$tobacco);

> covid\_dead\_test$contact\_other\_covid = as.factor(covid\_dead\_test$contact\_other\_covid);

> covid\_dead\_test$day\_cnt = as.numeric(covid\_dead\_test$day\_cnt);

> str(covid\_dead\_test);

'data.frame': 420 obs. of 16 variables:

$ sex : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

$ patient\_type : Factor w/ 1 level "2": 1 1 1 1 1 1 1 1 1 1 ...

$ pneumonia : Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 1 1 1 ...

$ age : num 66 68 49 61 47 56 68 72 78 64 ...

$ diabetes : Factor w/ 2 levels "1","2": 1 2 2 1 1 2 1 1 1 1 ...

$ copd : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 1 2 2 ...

$ asthma : Factor w/ 2 levels "1","2": 2 2 2 2 1 2 2 2 2 2 ...

$ inmsupr : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ hypertension : Factor w/ 2 levels "1","2": 1 2 2 1 2 2 1 1 2 2 ...

$ other\_disease : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ cardiovascular : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ obesity : Factor w/ 2 levels "1","2": 1 2 1 2 2 2 1 1 1 2 ...

$ renal\_chronic : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ tobacco : Factor w/ 2 levels "1","2": 2 2 2 2 2 2 2 2 2 2 ...

$ contact\_other\_covid: Factor w/ 2 levels "1","2": 1 2 2 2 2 2 2 2 2 2 ...

$ day\_cnt : num 3 5 16 4 18 10 9 6 6 8 ...

> head(covid\_dead\_test);

sex patient\_type pneumonia age diabetes copd asthma inmsupr hypertension other\_disease cardiovascular obesity renal\_chronic tobacco

1 1 2 1 66 1 2 2 2 1 2 2 1 2 2

2 1 2 1 68 2 2 2 2 2 2 2 2 2 2

3 1 2 1 49 2 2 2 2 2 2 2 1 2 2

4 1 2 1 61 1 2 2 2 1 2 2 2 2 2

5 1 2 1 47 1 2 1 2 2 2 2 2 2 2

6 1 2 1 56 2 2 2 2 2 2 2 2 2 2

contact\_other\_covid day\_cnt

1 1 3

2 2 5

3 2 16

4 2 4

5 2 18

6 2 10