Comprehensive Random Analysis of Covid-19

2023-01-24

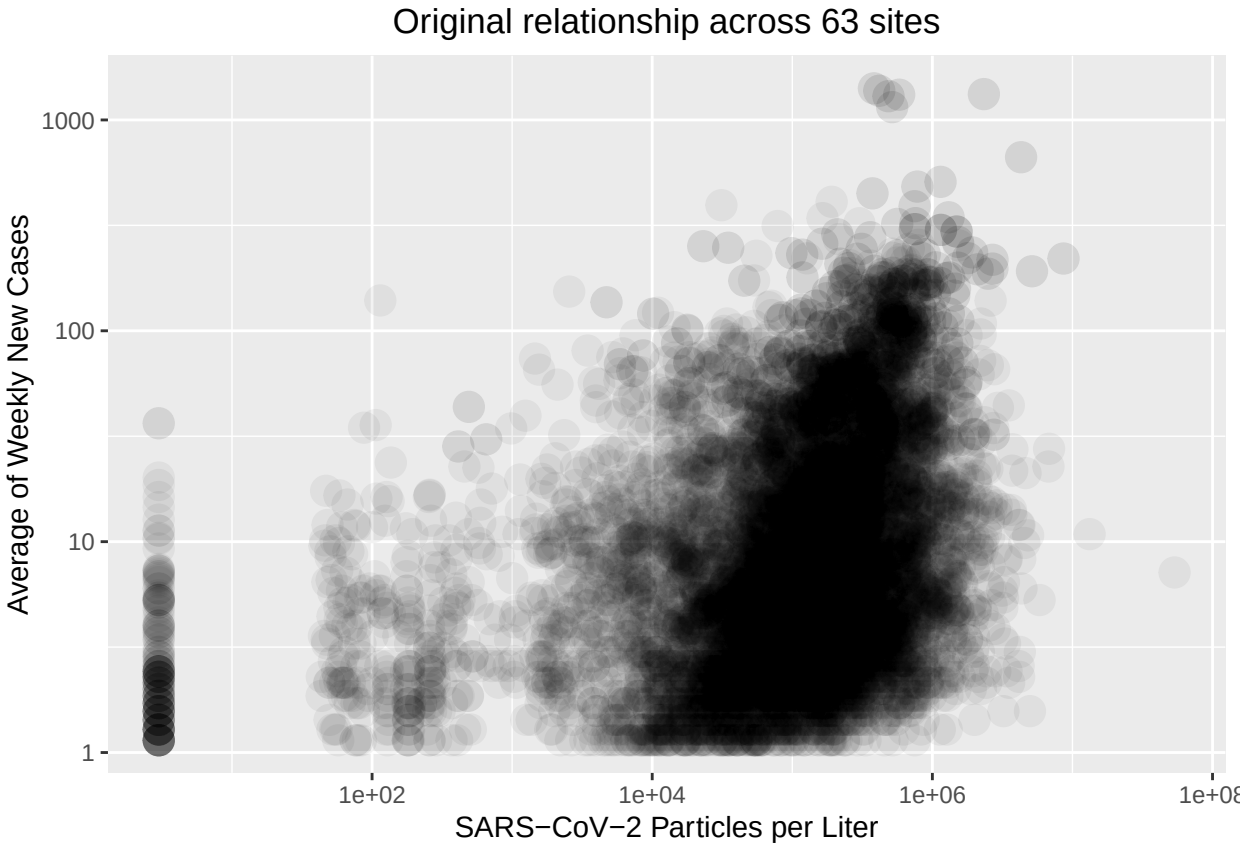
```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6110 -1.0416 -0.1479  0.9475  4.6796
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.11844    0.16972   0.698    0.485
## avg_sars_cov2_conc  0.19096    0.01373  13.906   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.276 on 6327 degrees of freedom
## Multiple R-squared:  0.02966,    Adjusted R-squared:  0.0295
## F-statistic: 193.4 on 1 and 6327 DF,  p-value: < 2.2e-16
```

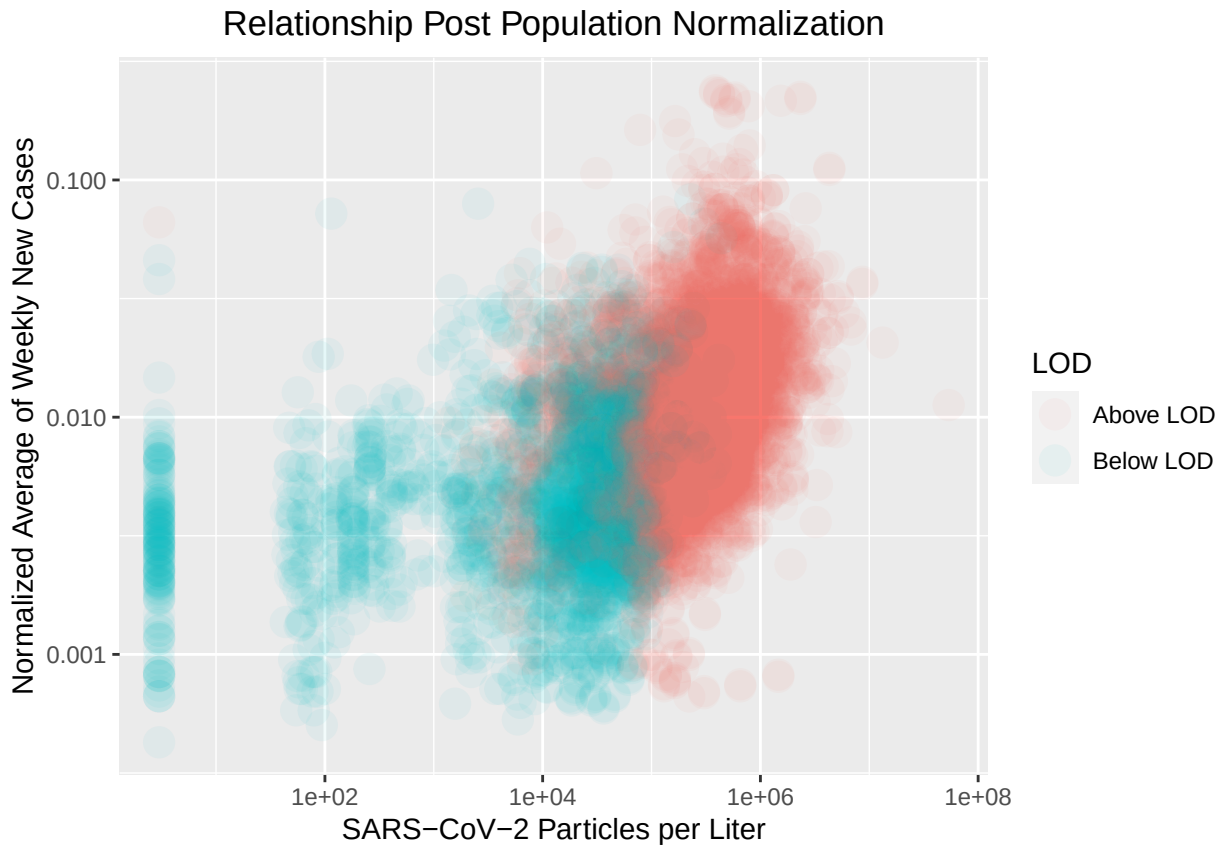
```
##
## Call:
## lm(formula = case_rate ~ flow_avg_conc, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1133 -0.5612 -0.0917  0.4737  6.0000
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -4.74411    0.07616  -62.29   <2e-16 ***
## flow_avg_conc  0.53052    0.00555   95.60   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8282 on 6327 degrees of freedom
## Multiple R-squared:  0.5909, Adjusted R-squared:  0.5908
## F-statistic:  9139 on 1 and 6327 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc:pop_group, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

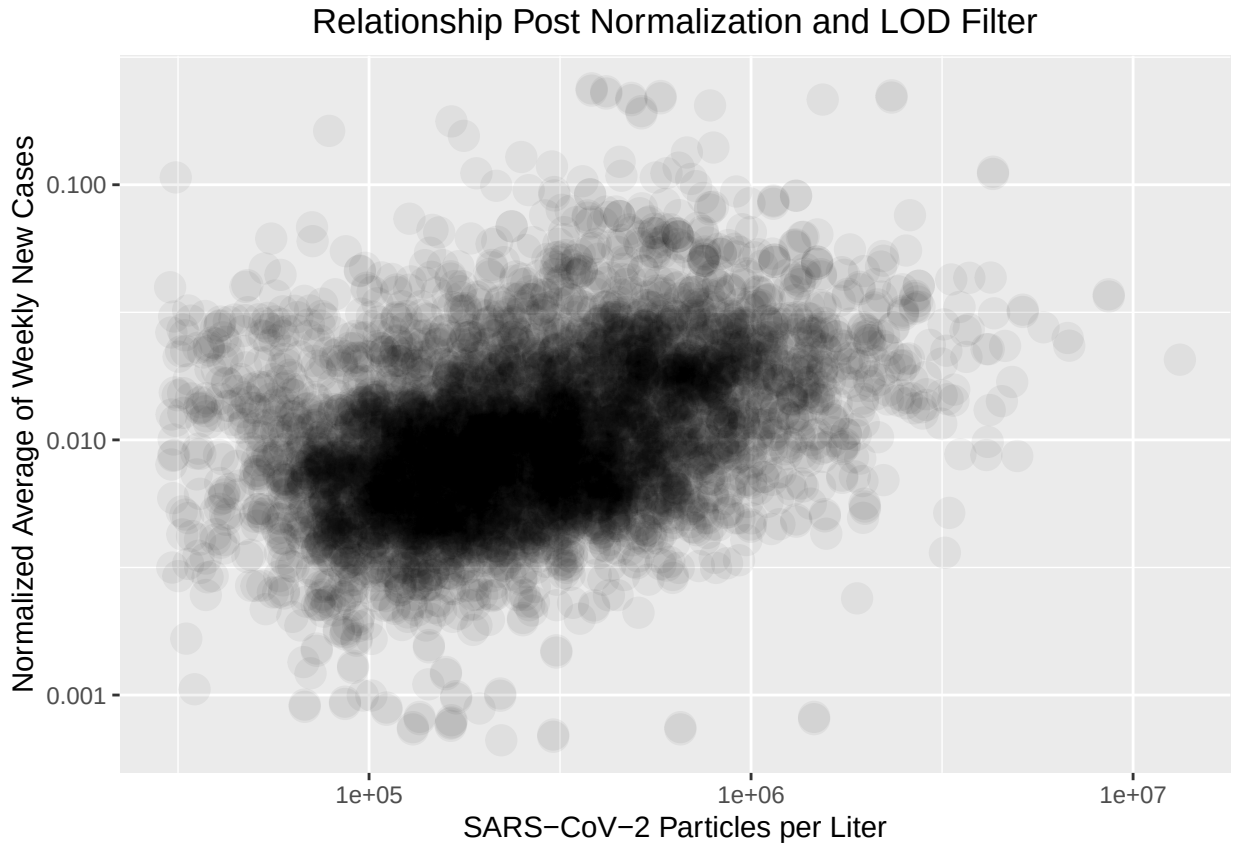
```
## -3.4692 -0.5295 -0.0220  0.4827  3.2351
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   -0.508338   0.103503  -4.911 9.27e-07 ***
## avg_sars_cov2_conc:pop_group1  0.141261   0.008346  16.926  < 2e-16 ***
## avg_sars_cov2_conc:pop_group2  0.191234   0.008546  22.377  < 2e-16 ***
## avg_sars_cov2_conc:pop_group3  0.271754   0.008576  31.687  < 2e-16 ***
## avg_sars_cov2_conc:pop_group4  0.352087   0.008482  41.511  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7734 on 6324 degrees of freedom
## Multiple R-squared:  0.6434, Adjusted R-squared:  0.6432
## F-statistic:  2853 on 4 and 6324 DF,  p-value: < 2.2e-16


##
## Call:
## lm(formula = case_rate ~ flow_avg_conc:pop_group, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0409 -0.4944 -0.0497  0.4215  3.2404
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -1.681951   0.098044  -17.16   <2e-16 ***
## flow_avg_conc:pop_group1  0.246145   0.008263   29.79   <2e-16 ***
## flow_avg_conc:pop_group2  0.275346   0.007773   35.42   <2e-16 ***
## flow_avg_conc:pop_group3  0.322068   0.007113   45.28   <2e-16 ***
## flow_avg_conc:pop_group4  0.354316   0.006362   55.69   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.728 on 6324 degrees of freedom
## Multiple R-squared:  0.6841, Adjusted R-squared:  0.6839
## F-statistic:  3423 on 4 and 6324 DF,  p-value: < 2.2e-16
```


Original relationship across 63 sites
Average of Weekly New Cases
1000
100
10
1
1e+02
1e+04
1e+06
SARS−CoV−2 Particles per Liter



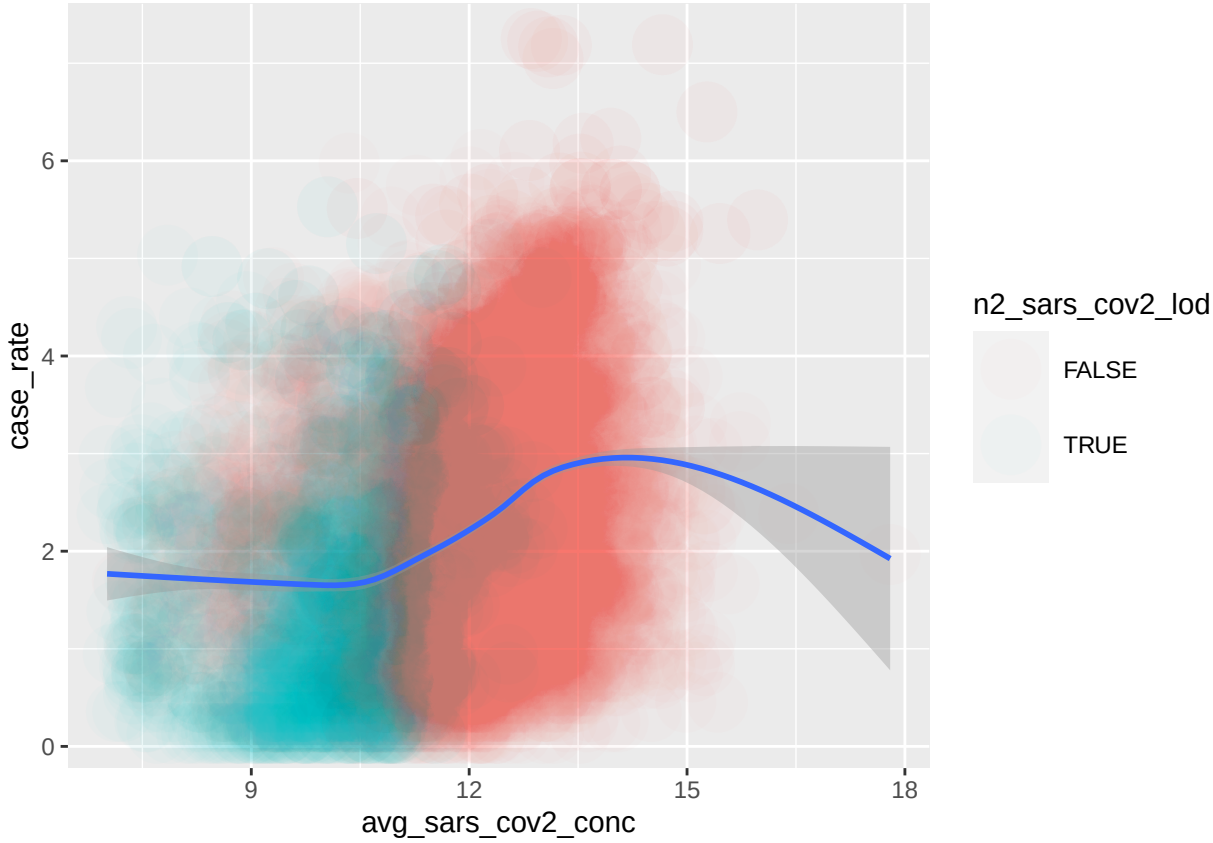

```
##
## Call:
## lm(formula = n_case_rate ~ avg_sars_cov2_conc, data = .)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9869 -0.4610 -0.0356  0.3880  3.2453
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -9.48503    0.17131  -55.37   <2e-16 ***
## avg_sars_cov2_conc  0.39240    0.01378   28.47   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6965 on 5284 degrees of freedom
## Multiple R-squared:  0.133,  Adjusted R-squared:  0.1328
## F-statistic: 810.6 on 1 and 5284 DF,  p-value: < 2.2e-16
```

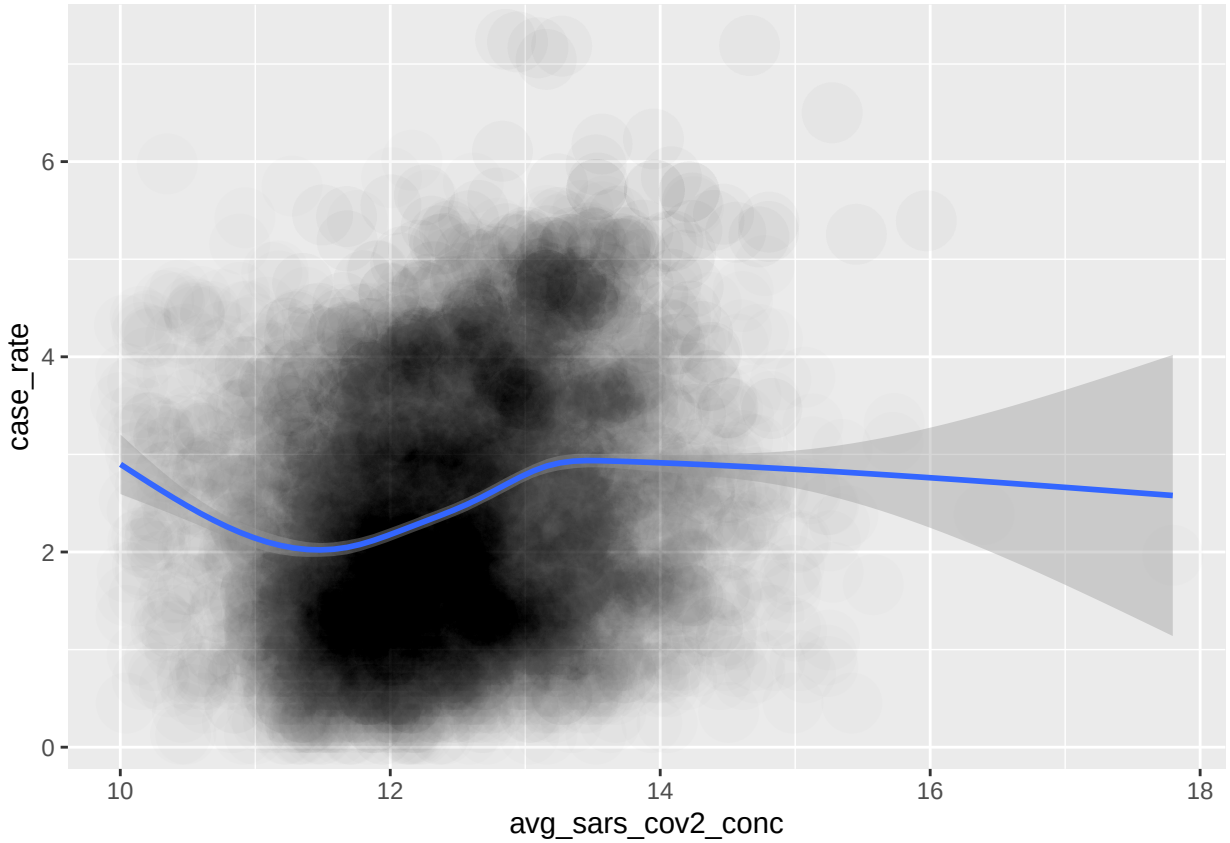


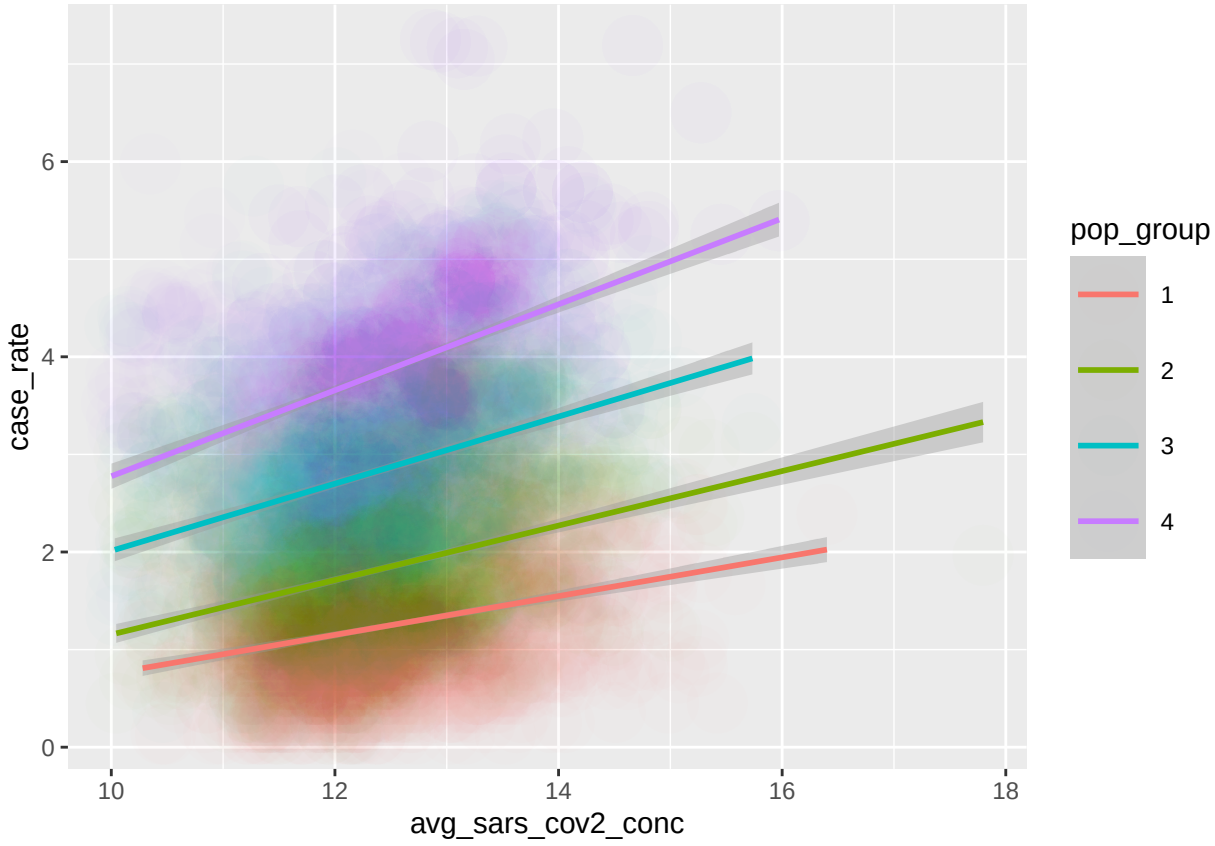

```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc, data = .)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4786 -1.0003 -0.1675  0.8605  4.7966
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.079002   0.065870   1.199     0.23
## avg_sars_cov2_conc 0.184926   0.005678  32.571   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.233 on 9012 degrees of freedom
## Multiple R-squared:  0.1053, Adjusted R-squared:  0.1052
## F-statistic:  1061 on 1 and 9012 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc + log(pop), data = .)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -3.10181 -0.47451 -0.03898  0.40102  3.12664
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -8.496980   0.140623  -60.42   <2e-16 ***
## avg_sars_cov2_conc  0.313553   0.010163   30.85   <2e-16 ***
## log(pop)            0.677734   0.005844  115.98   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7096 on 5971 degrees of freedom
## Multiple R-squared:  0.7076, Adjusted R-squared:  0.7075
## F-statistic:  7226 on 2 and 5971 DF,  p-value: < 2.2e-16
```




case_rate
6
4
2
0
10
12
14
16
18
avg_sars_cov2_conc

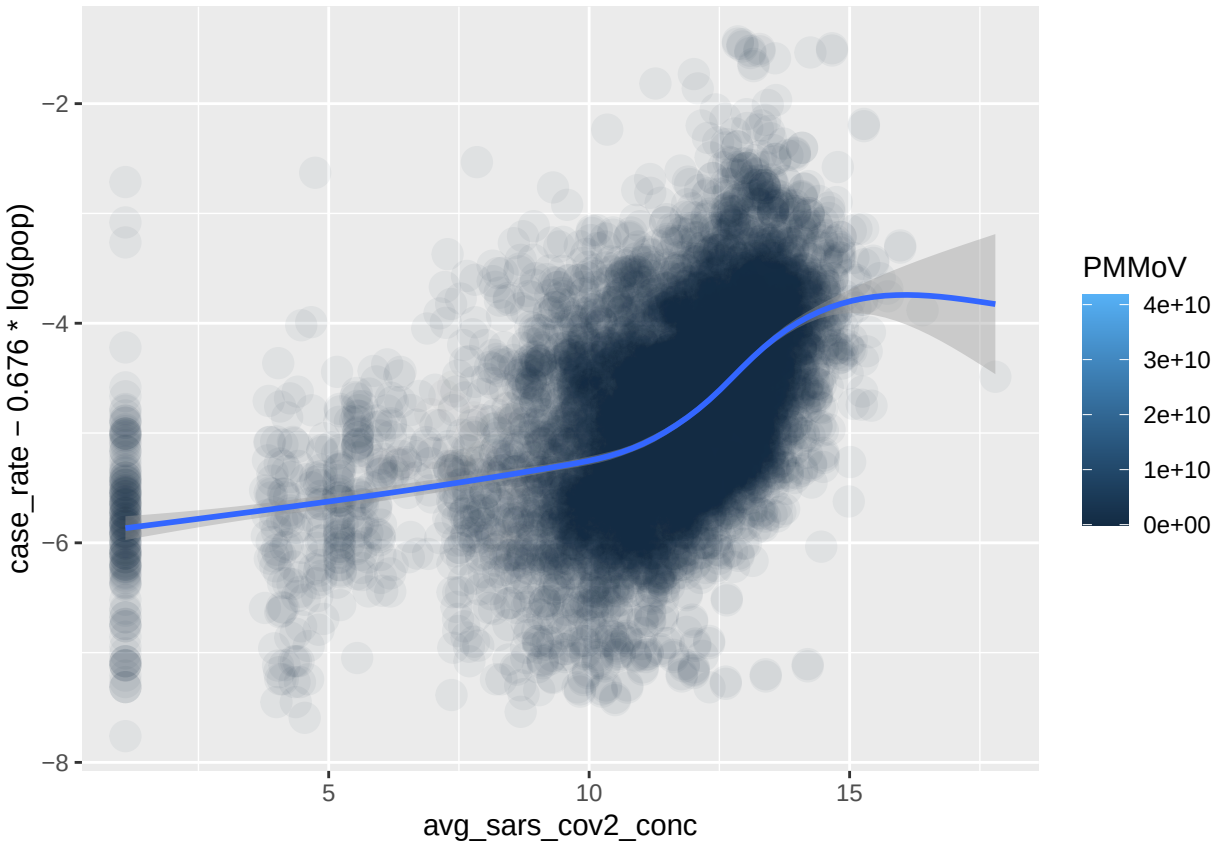


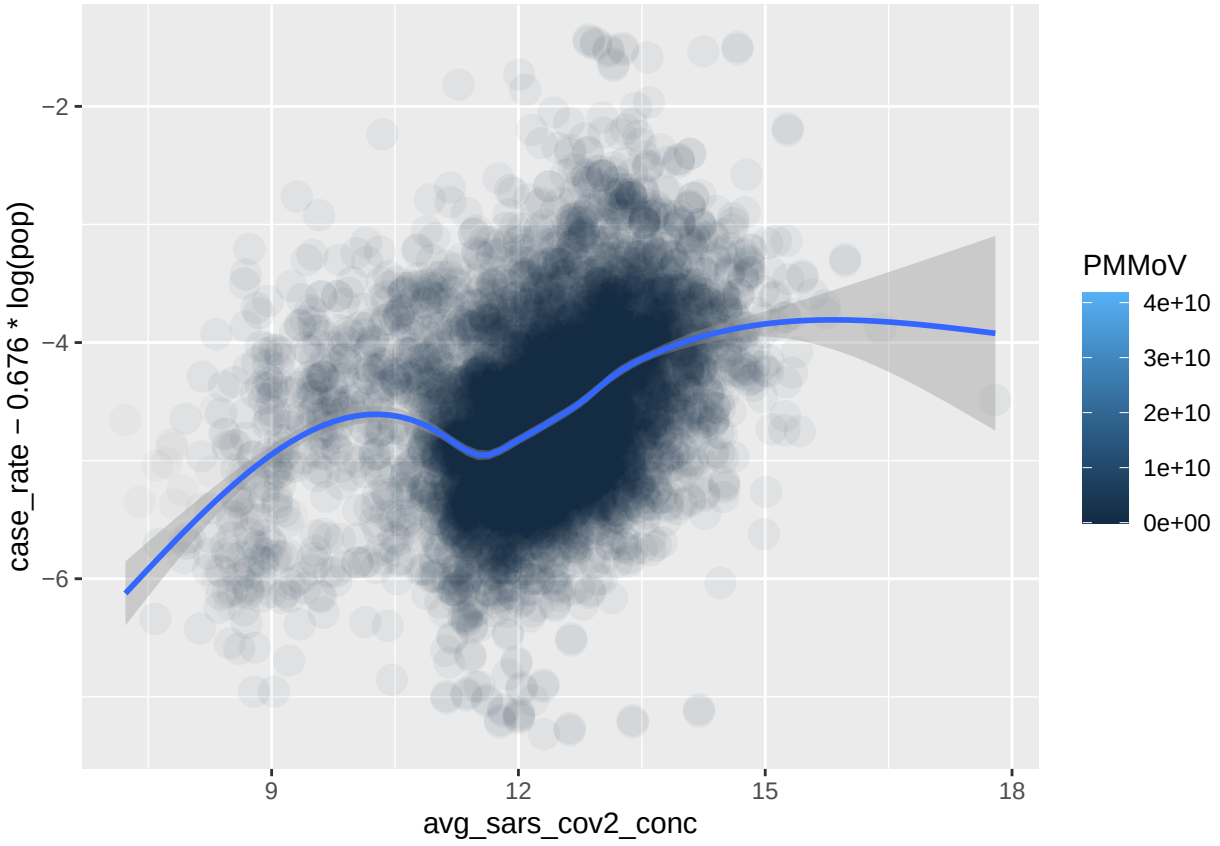

```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc, data = Graph_DF)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4786 -1.0003 -0.1675  0.8605  4.7966
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.079002   0.065870   1.199     0.23
## avg_sars_cov2_conc 0.184926   0.005678  32.571   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.233 on 9012 degrees of freedom
## Multiple R-squared:  0.1053, Adjusted R-squared:  0.1052
## F-statistic:  1061 on 1 and 9012 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -2.6110 -1.0416 -0.1479  0.9475  4.6796
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.11844    0.16972   0.698    0.485
## avg_sars_cov2_conc  0.19096    0.01373  13.906   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.276 on 6327 degrees of freedom
## Multiple R-squared:  0.02966,    Adjusted R-squared:  0.0295
## F-statistic: 193.4 on 1 and 6327 DF,  p-value: < 2.2e-16


##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc + pop_group, data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4628 -0.5369 -0.0232  0.4875  3.2967
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -1.886231   0.107872  -17.49   <2e-16 ***
## avg_sars_cov2_conc  0.250857   0.008455   29.67   <2e-16 ***
## pop_group2          0.654168   0.028213   23.19   <2e-16 ***
## pop_group3          1.635104   0.028399   57.58   <2e-16 ***
## pop_group4          2.617573   0.027218   96.17   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7796 on 6324 degrees of freedom
## Multiple R-squared:  0.6377, Adjusted R-squared:  0.6374
## F-statistic:  2782 on 4 and 6324 DF,  p-value: < 2.2e-16
```


case_rate − 0.676 * log(pop)
−2
−4
−6
−8
5
10
15
avg_sars_cov2_conc
PMMoV
4e+10
3e+10
2e+10
1e+10
0e+00

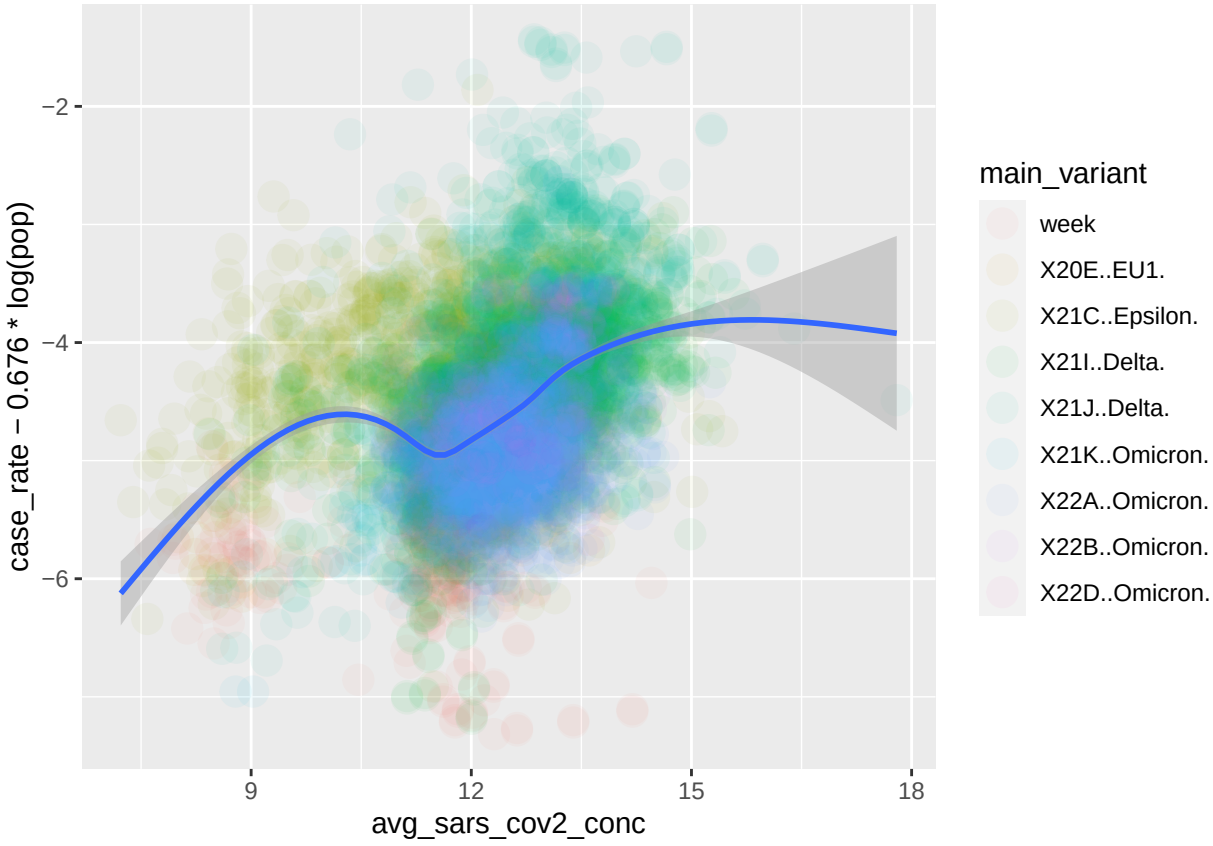



```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc + log(pop), data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9662 -0.4952 -0.0512  0.4306  4.3759
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.335831   0.116258  -63.10   <2e-16 ***
## avg_sars_cov2_conc  0.223741   0.007815   28.63   <2e-16 ***
## log(pop)            0.676036   0.005876  115.04   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7254 on 6326 degrees of freedom
## Multiple R-squared:  0.6862, Adjusted R-squared:  0.6861
## F-statistic:  6916 on 2 and 6326 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc, data = .)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -2.6110 -1.0416 -0.1479  0.9475  4.6796
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.11844    0.16972   0.698    0.485
## avg_sars_cov2_conc   0.19096    0.01373  13.906   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.276 on 6327 degrees of freedom
## Multiple R-squared:  0.02966,    Adjusted R-squared:  0.0295
## F-statistic: 193.4 on 1 and 6327 DF,  p-value: < 2.2e-16


##
## Call:
## lm(formula = n_case_rate ~ avg_sars_cov2_conc, data = mod_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9661 -0.4951 -0.0512  0.4306  4.3757
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.335432   0.096519  -76.00   <2e-16 ***
## avg_sars_cov2_conc  0.223739   0.007809   28.65   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7254 on 6327 degrees of freedom
## Multiple R-squared:  0.1148, Adjusted R-squared:  0.1147
## F-statistic: 820.8 on 1 and 6327 DF,  p-value: < 2.2e-16
```




```
##
## Call:
## lm(formula = n_case_rate ~ avg_sars_cov2_conc, data = rm_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9661 -0.4951 -0.0512  0.4306  4.3757
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.335432   0.096519  -76.00   <2e-16 ***
## avg_sars_cov2_conc  0.223739   0.007809   28.65   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7254 on 6327 degrees of freedom
## Multiple R-squared:  0.1148, Adjusted R-squared:  0.1147
## F-statistic: 820.8 on 1 and 6327 DF,  p-value: < 2.2e-16
```

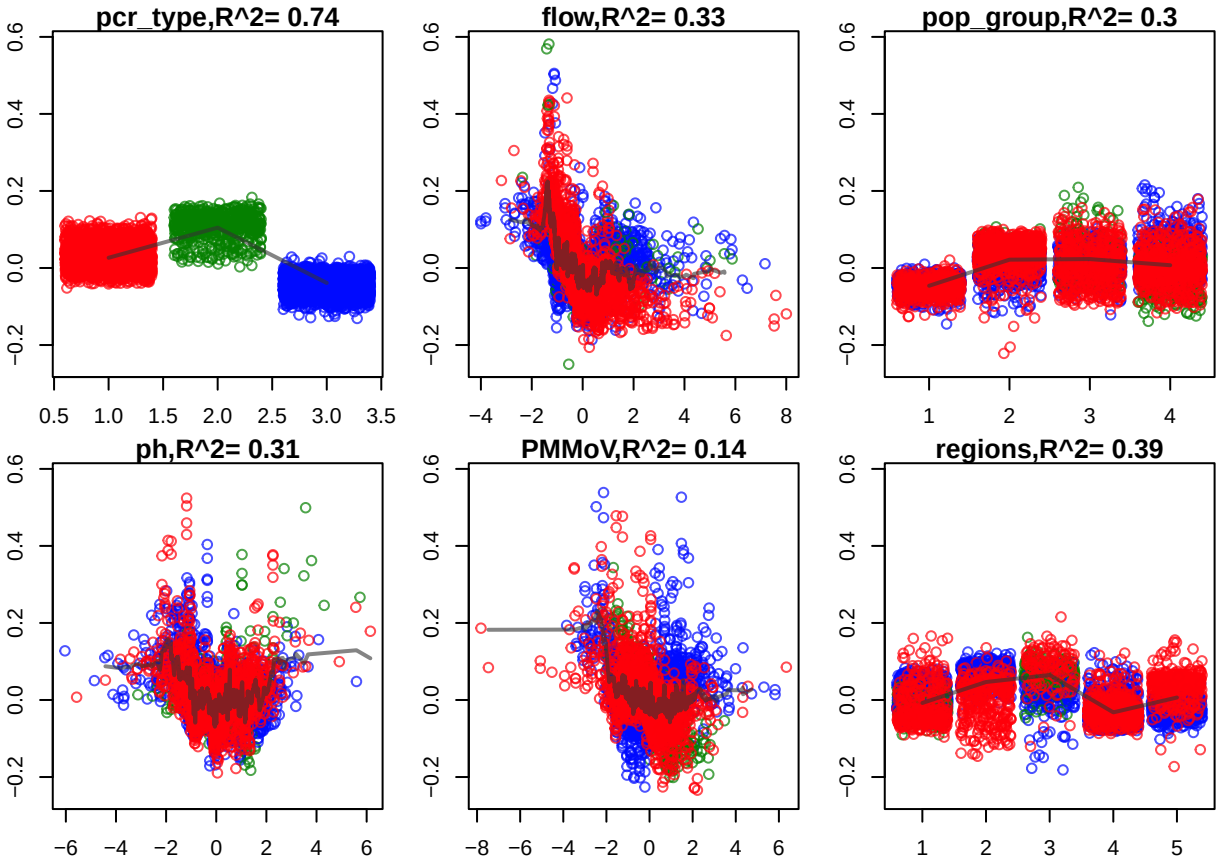
```
##
## Call:
##  randomForest(formula = resid ~ ., data = rm_data, ntree = 500,      importance = TRUE, keep.inbag =
##                Type of random forest: regression
##                      Number of trees: 500
## No. of variables tried at each split: 2
```

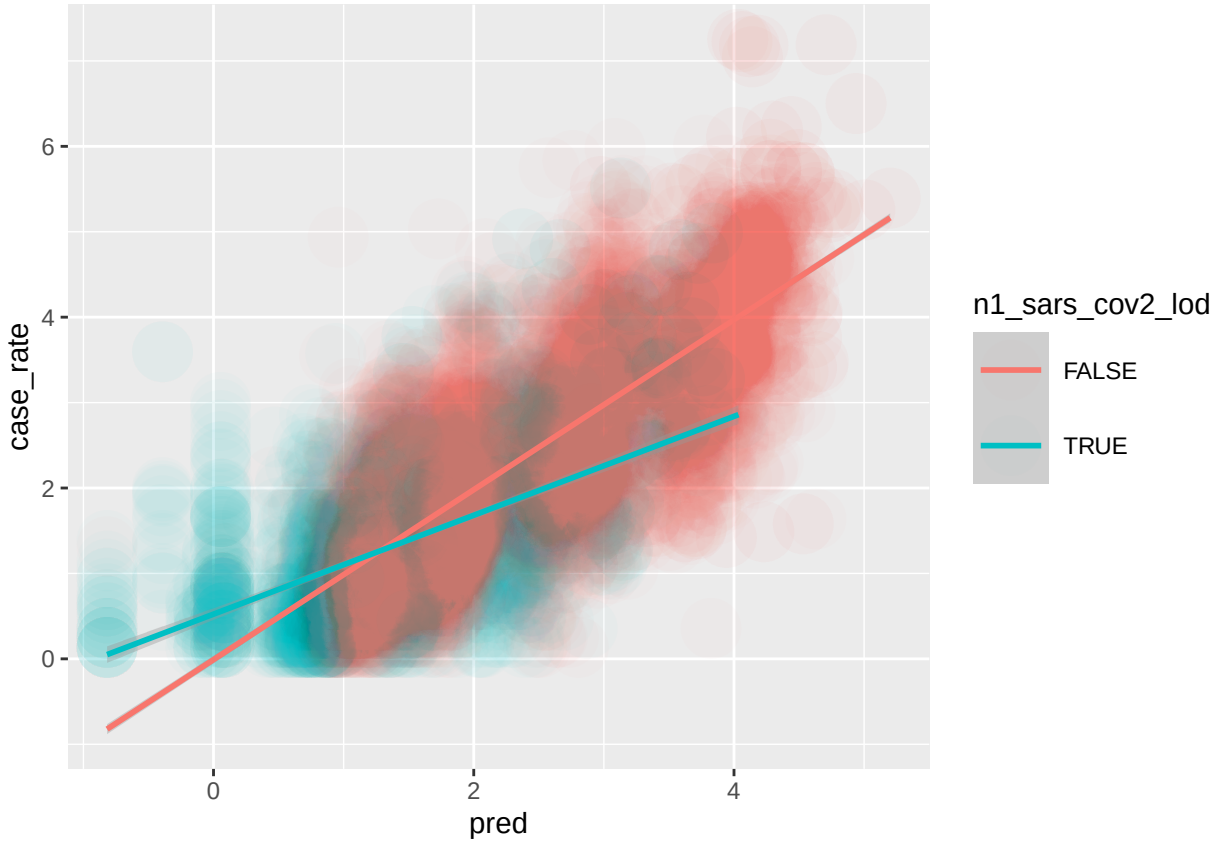
	%IncMSE	IncNodePurity
regions	63.10622	73.42110
PMMoV	68.07717	274.73034
ph	72.08750	230.92665
pop_group	75.92171	70.63900
flow	80.73994	272.57386
pcr_type	80.96885	51.43222

```
##
##           Mean of squared residuals: 0.1483788
##                     % Var explained: 28.86

##             %IncMSE IncNodePurity
## pcr_type  80.96885       51.43222
## pop_group 75.92171       70.63900
## regions   63.10622       73.42110
## ph        72.08751      230.92665
## flow      80.73994      272.57386
## PMMoV     68.07717      274.73034
```

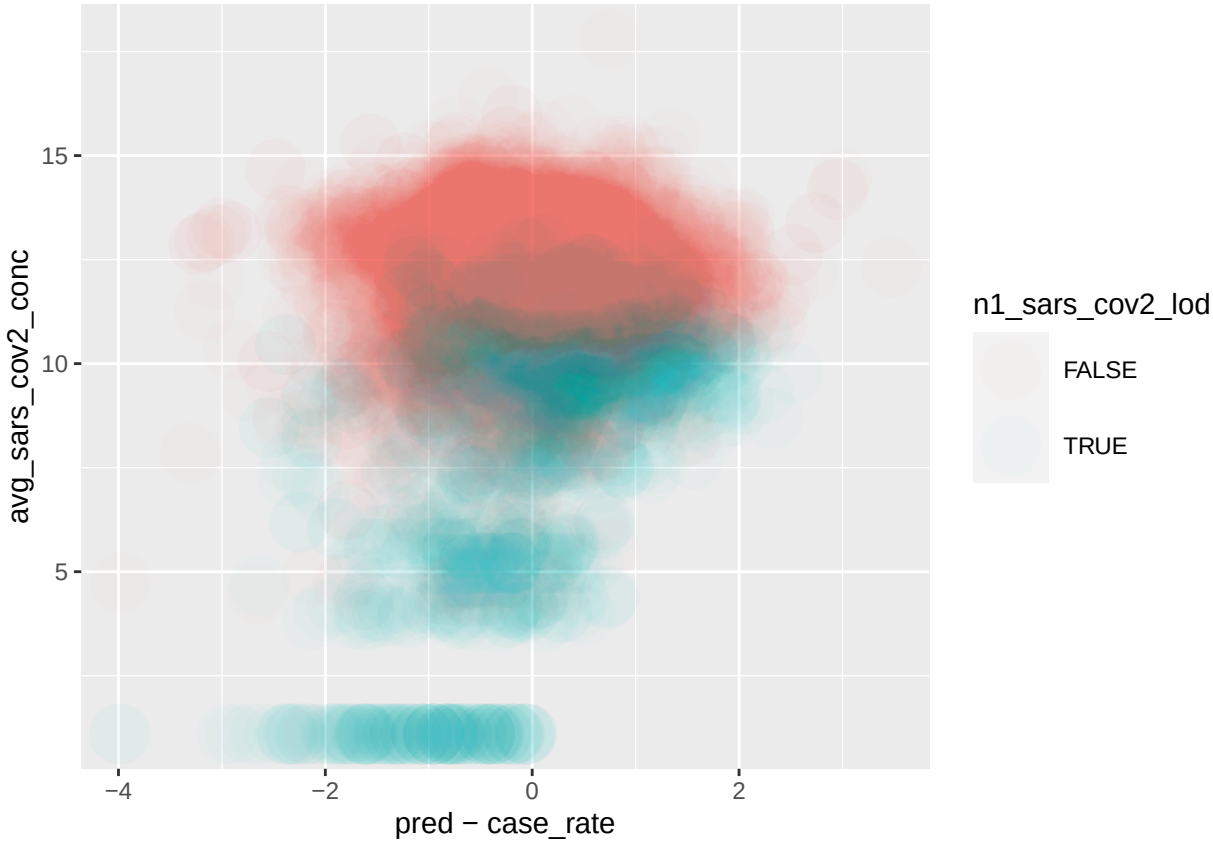
```
## [1] "compute goodness-of-fit with leave-one-out k-nearest neighbor(guassian weighting), kknn package"
```






```
##
## Call:
## lm(formula = case_rate ~ avg_sars_cov2_conc:pop_group + pop_group,
##     data = df_LOD)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4671 -0.5307 -0.0177  0.4799  3.2242
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    -1.01763    0.24530  -4.148 3.39e-05 ***
## pop_group2                      0.92297    0.31456   2.934  0.00336 **
## pop_group3                      0.79942    0.32483   2.461  0.01388 *
## pop_group4                      0.21299    0.30760   0.692  0.48869
## avg_sars_cov2_conc:pop_group1   0.18161    0.01949   9.316  < 2e-16 ***
## avg_sars_cov2_conc:pop_group2   0.15769    0.01605   9.826  < 2e-16 ***
## avg_sars_cov2_conc:pop_group3   0.24816    0.01740  14.265  < 2e-16 ***
## avg_sars_cov2_conc:pop_group4   0.37599    0.01504  24.993  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7728 on 6321 degrees of freedom
## Multiple R-squared:  0.6442, Adjusted R-squared:  0.6438
## F-statistic:  1635 on 7 and 6321 DF,  p-value: < 2.2e-16
```

	adjusted R^2	mse	num_param
All categorical interaction model	0.771	0.383	145
All categorical indirect model	0.756	0.413	49
Sub data all interaction model	0.767	0.386	73
Sub data indirect interaction model	0.755	0.410	24
Base relationship	0.139	1.464	3



////////////////////////////////////// // HFG Work //////////////////////////////////////