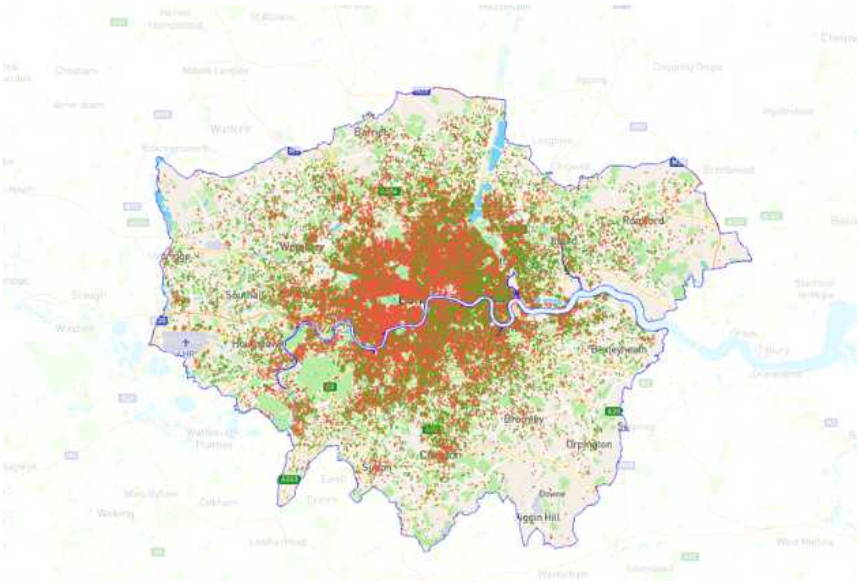
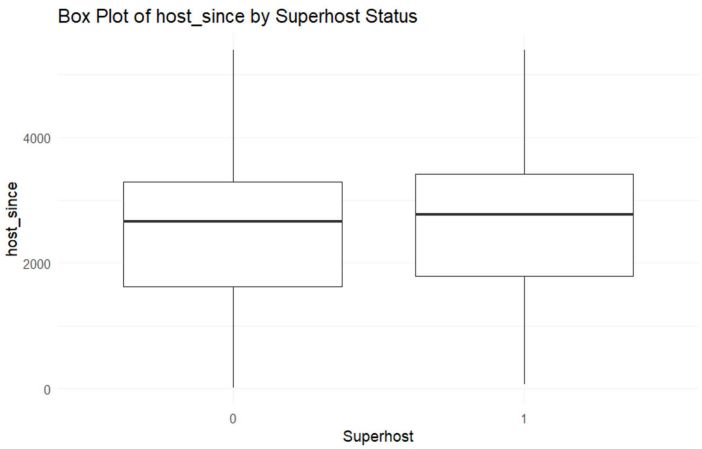
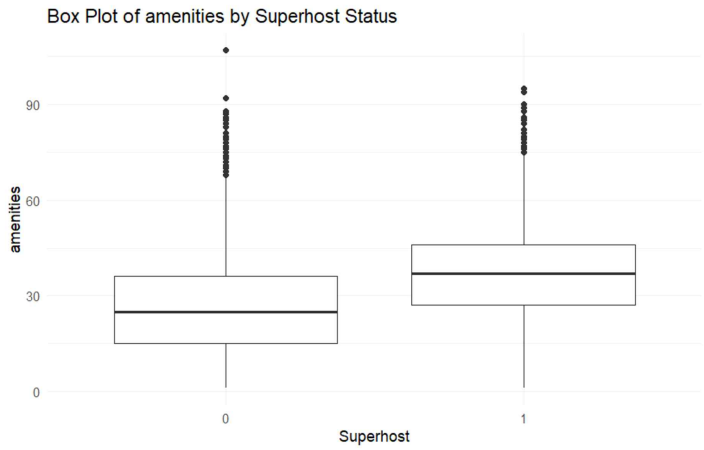
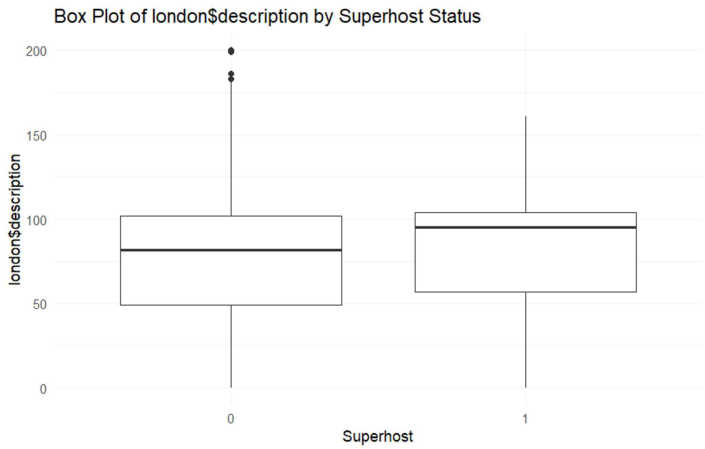
부록 1

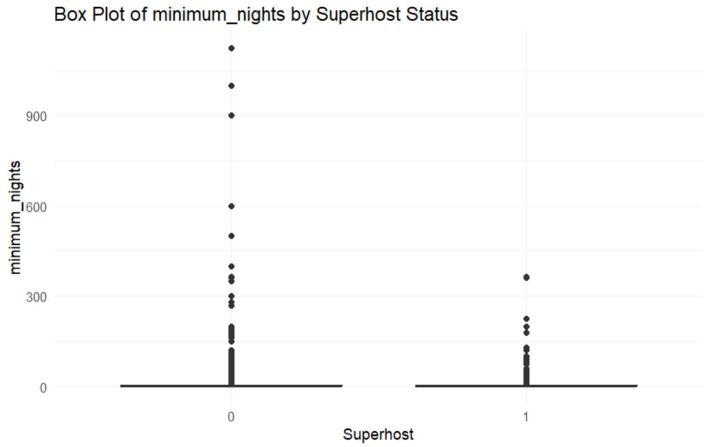


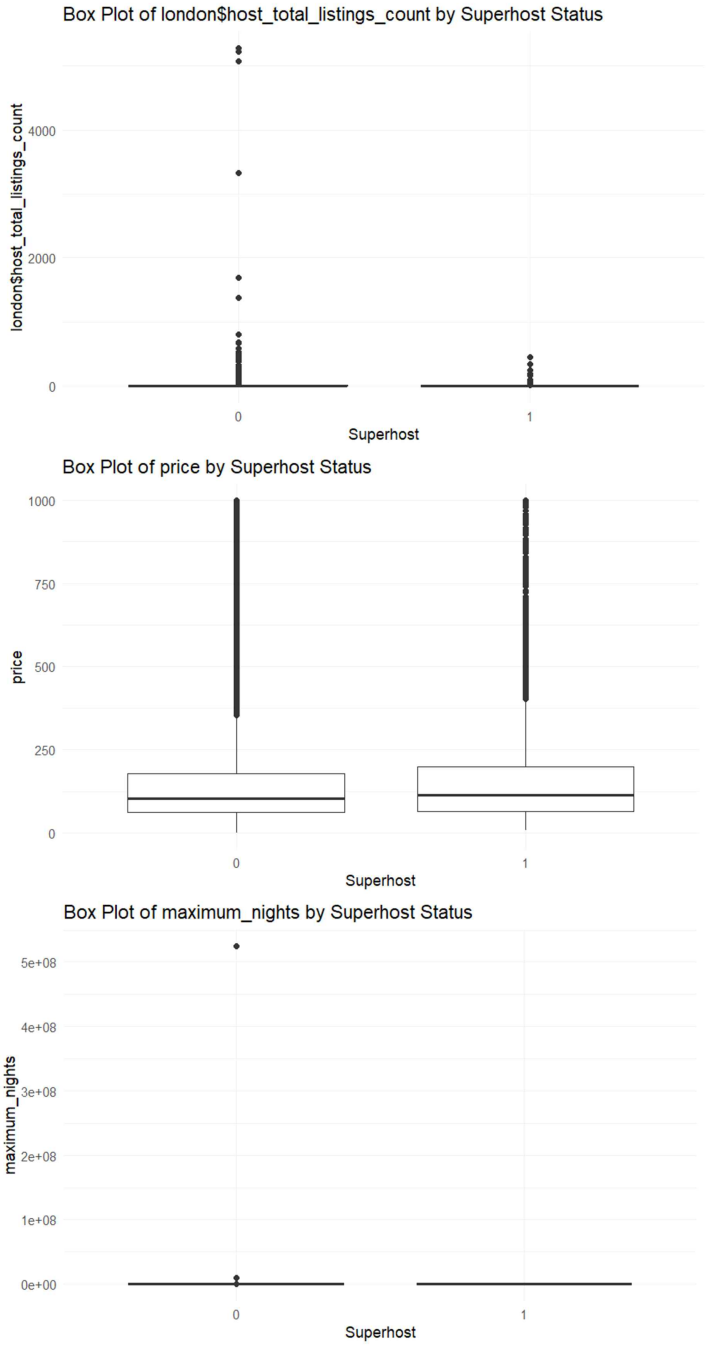
<그림 1>


Box Plot of host_since by Superhost Status
host_since
4000
2000
0
0
1
Superhost

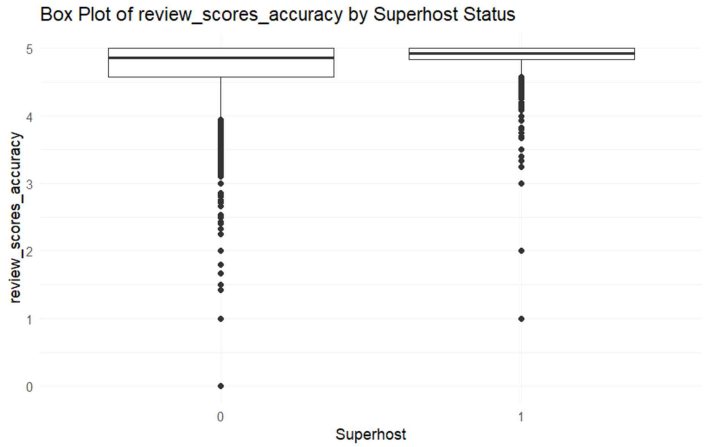

Box Plot of amenities by Superhost Status
amenities
90
60
30
0
0
1
Superhost

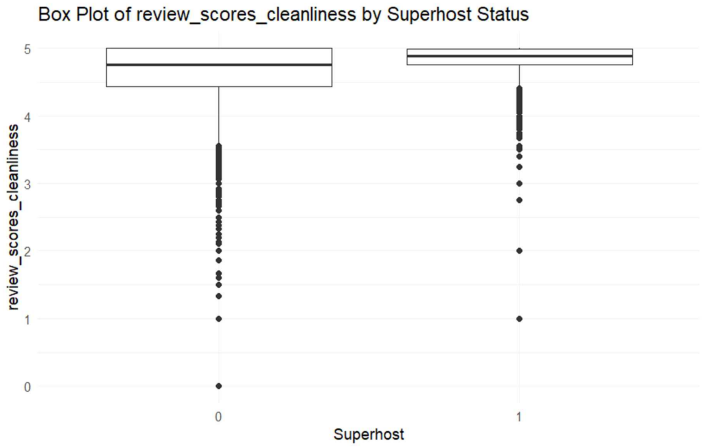

Box Plot of london$description by Superhost Status
200
150
100
50
0
london$description
0
1
Superhost

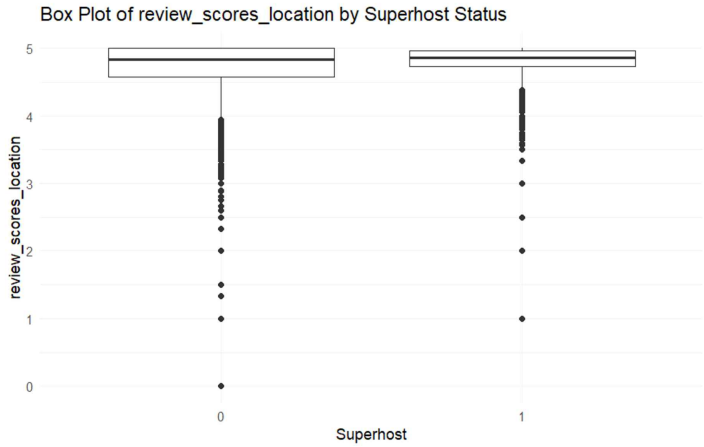

Box Plot of minimum_nights by Superhost Status
900
600
300
0
minimum_nights
0
1
Superhost

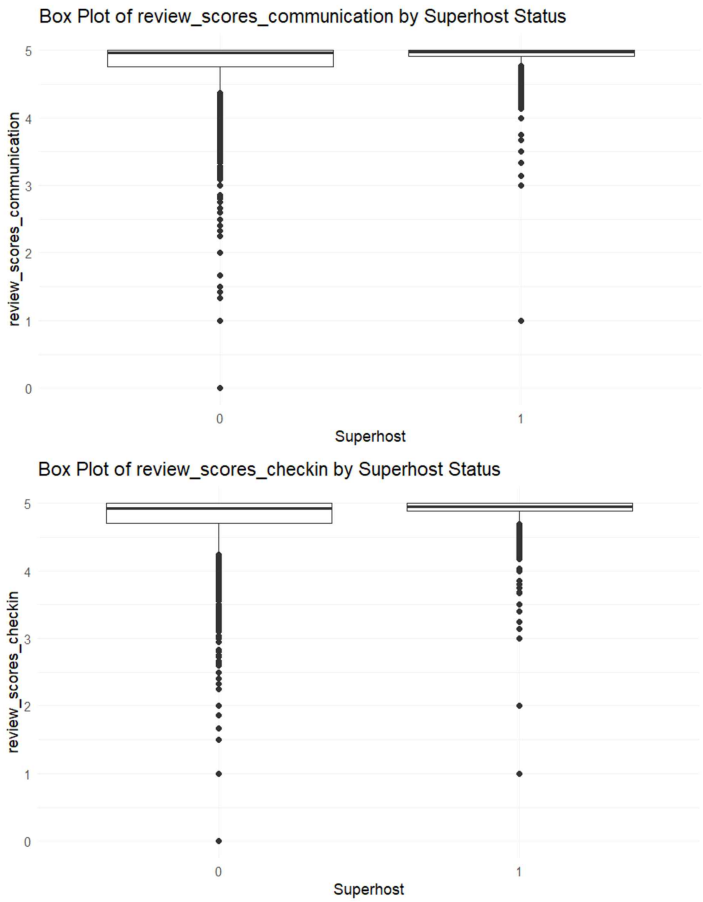

Box Plot of london$host_total_listings_count by Superhost Status
london$host_total_listings_count
4000
2000
0
0
1
Superhost
Box Plot of price by Superhost Status
price
1000
750
500
250
0
0
1
Superhost
Box Plot of maximum_nights by Superhost Status
maximum_nights
5e+08
4e+08
3e+08
2e+08
1e+08
0e+00
0
1
Superhost

<그림 2>


Box Plot of review_scores_accuracy by Superhost Status
review_scores_accuracy
0
1
2
3
4
5
0
1
Superhost


Box Plot of review_scores_cleanliness by Superhost Status
review_scores_cleanliness
0
1
2
3
4
5
0
1
Superhost


Box Plot of review_scores_location by Superhost Status
review_scores_location
0
1
2
3
4
5
0
1
Superhost


Box Plot of review_scores_communication by Superhost Status
review_scores_communication
Superhost
Box Plot of review_scores_checkin by Superhost Status
review_scores_checkin
Superhost

<그림 3>

부록 2

```
> deviance_value <- anova(mod2, test = "Chisq") # deviance 확인
> deviance_value
Analysis of Deviance Table

Model: binomial, link: logit

Response: host_is_superhost

Terms added sequentially (first to last)


                          Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                                      45163      41164
description                1    736.4     45162      40427 < 2.2e-16 ***
host_since                 1     21.2     45161      40406 4.093e-06 ***
host_about                 1    200.7     45160      40205 < 2.2e-16 ***
host_response_time         4   5596.1     45156      34609 < 2.2e-16 ***
host_response_rate         1    380.2     45155      34229 < 2.2e-16 ***
host_total_listings_count  1    265.7     45154      33963 < 2.2e-16 ***
host_has_profile_pic       1     23.3     45153      33940 1.353e-06 ***
room_type                  3    409.6     45150      33531 < 2.2e-16 ***
accommodates               1     25.0     45149      33505 5.606e-07 ***
bathrooms_text             1      5.0     45148      33500   0.02474 *
amenities                  1    918.9     45147      32582 < 2.2e-16 ***
price                      1    228.5     45146      32353 < 2.2e-16 ***
minimum_nights             1     21.1     45145      32332 4.423e-06 ***
maximum_nights             1      0.0     45144      32332   0.92866
number_of_reviews_ltm      1   1417.1     45143      30915 < 2.2e-16 ***
instant_bookable           1     85.5     45142      30829 < 2.2e-16 ***
review_scores              1   3903.1     45141      26926 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

<그림1>

```
> summary(mod0) # AIC: 24037

Call:
glm(formula = host_is_superhost ~ ., family = binomial(link = "logit"),
    data = london)

Coefficients:
                                    Estimate Std. Error z value Pr(>|z|)
(Intercept)                       -5.084e+01  1.165e+00 -43.625  < 2e-16 ***
description                        4.209e-03  5.714e-04   7.366 1.76e-13 ***
host_since                         1.319e-04  1.495e-05   8.821  < 2e-16 ***
host_about                         5.187e-03  6.001e-04   8.644  < 2e-16 ***
host_response_timeNoInfo           7.006e-01  2.881e-01   2.431 0.015038 *
host_response_timewithin a day    -7.047e-01  3.531e-01  -1.996 0.045958 *
host_response_timewithin a few hours -4.167e-01  3.605e-01  -1.156 0.247635
host_response_timewithin an hour  -1.126e-01  3.619e-01  -0.311 0.755735
host_response_rate                 3.139e-02  3.204e-03   9.797  < 2e-16 ***
host_total_listings_count          8.158e-03  4.317e-03   1.890 0.058793 .
host_has_profile_pic1              2.876e-01  1.198e-01   2.400 0.016409 *
room_typeHotel room               -1.221e+00  8.715e-01  -1.401 0.161299
room_typePrivate room              5.012e-01  4.688e-02  10.692  < 2e-16 ***
room_typeShared room              -6.270e-01  3.689e-01  -1.699 0.089254 .
accommodates                      -1.150e-01  1.470e-02  -7.823 5.16e-15 ***
bathrooms_text                    -1.836e-01  4.007e-02  -4.581 4.64e-06 ***
amenities                          2.013e-02  1.213e-03  16.598  < 2e-16 ***
price                              3.309e-03  3.581e-04   9.241  < 2e-16 ***
minimum_nights                     9.079e-02  1.368e-02   6.639 3.15e-11 ***
maximum_nights                     1.523e-04  3.460e-05   4.403 1.07e-05 ***
number_of_reviews_ltm              5.532e-02  1.322e-03  41.847  < 2e-16 ***
review_scores_rating               7.031e+00  2.221e-01  31.652  < 2e-16 ***
review_scores_accuracy             5.311e-01  1.582e-01   3.358 0.000785 ***
review_scores_cleanliness          6.185e-01  1.001e-01   6.182 6.33e-10 ***
review_scores_checkin              2.318e-01  1.300e-01   1.783 0.074555 .
review_scores_communication        1.086e+00  2.177e-01   4.988 6.10e-07 ***
review_scores_location            -4.506e-01  7.779e-02  -5.792 6.94e-09 ***
instant_bookable1                 -1.582e-01  4.280e-02  -3.696 0.000219 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 38590  on 42879  degrees of freedom
Residual deviance: 23981  on 42852  degrees of freedom
AIC: 24037

Number of Fisher Scoring iterations: 8
```

<그림2>

```
> summary(mod1) # AIC: 25910

Call:
glm(formula = host_is_superhost ~ ., family = binomial(link = "logit"),
    data = london)

Coefficients:
                                  Estimate Std. Error z value Pr(>|z|)
(Intercept)                     -5.215e+01  1.124e+00 -46.380  < 2e-16 ***
description                      4.116e-03  5.502e-04   7.480 7.44e-14 ***
host_since                       1.343e-04  1.431e-05   9.382  < 2e-16 ***
host_about                       5.247e-03  5.719e-04   9.176  < 2e-16 ***
host_response_timeNoInfo         7.376e-01  2.787e-01   2.646 0.008139 **
host_response_timewithin a day  -7.916e-01  3.413e-01  -2.320 0.020358 *
host_response_timewithin a few hours -5.040e-01  3.484e-01  -1.447 0.148018
host_response_timewithin an hour -2.118e-01  3.498e-01  -0.606 0.544786
host_response_rate               3.233e-02  3.086e-03  10.475  < 2e-16 ***
host_total_listings_count        7.690e-03  4.066e-03   1.891 0.058610 .
host_has_profile_pic1            3.131e-01  1.169e-01   2.677 0.007423 **
room_typeHotel room             -1.605e+00  8.393e-01  -1.912 0.055905 .
room_typePrivate room            5.548e-01  4.787e-02  11.592  < 2e-16 ***
room_typeShared room            -5.341e-01  3.708e-01  -1.440 0.149760
accommodates                    -1.513e-01  1.524e-02  -9.928  < 2e-16 ***
bathrooms_text                   6.288e-02  1.919e-02   3.277 0.001049 **
amenities                        1.976e-02  1.157e-03  17.081  < 2e-16 ***
price                            3.894e-01  3.815e-02  10.208  < 2e-16 ***
minimum_nights                   8.628e-02  1.307e-02   6.603 4.02e-11 ***
maximum_nights                   1.754e-04  3.334e-05   5.261 1.44e-07 ***
number_of_reviews_ltm            5.531e-02  1.286e-03  42.999  < 2e-16 ***
review_scores_rating             6.932e+00  2.136e-01  32.455  < 2e-16 ***
review_scores_accuracy           5.201e-01  1.508e-01   3.448 0.000565 ***
review_scores_cleanliness        6.302e-01  9.705e-02   6.494 8.36e-11 ***
review_scores_checkin            2.627e-01  1.252e-01   2.099 0.035803 *
review_scores_communication      1.041e+00  2.075e-01   5.015 5.30e-07 ***
review_scores_location          -4.178e-01  7.652e-02  -5.460 4.77e-08 ***
instant_bookable1               -1.490e-01  4.099e-02  -3.634 0.000279 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 41164  on 45163  degrees of freedom
Residual deviance: 25854  on 45136  degrees of freedom
AIC: 25910

Number of Fisher Scoring iterations: 8
```

<그림3>

```
> summary(mod2) # AIC: 26972

Call:
glm(formula = host_is_superhost ~ ., family = binomial(link = "logit"),
    data = london)

Coefficients:
                                  Estimate Std. Error z value Pr(>|z|)
(Intercept)                     -4.869e+01  9.317e-01 -52.261  < 2e-16 ***
description                      3.309e-03  5.402e-04   6.125 9.09e-10 ***
host_since                       1.240e-04  1.405e-05   8.824  < 2e-16 ***
host_about                       4.868e-03  5.504e-04   8.843  < 2e-16 ***
host_response_timeNoInfo         6.843e-01  2.773e-01   2.468 0.013595 *
host_response_timewithin a day  -7.471e-01  3.385e-01  -2.207 0.027318 *
host_response_timewithin a few hours -4.560e-01  3.456e-01  -1.319 0.187006
host_response_timewithin an hour -1.686e-01  3.469e-01  -0.486 0.626889
host_response_rate               3.169e-02  3.044e-03  10.410  < 2e-16 ***
host_total_listings_count       -6.147e-03  3.948e-03  -1.557 0.119541
host_has_profile_pic1            2.957e-01  1.159e-01   2.550 0.010758 *
room_typeHotel room             -1.689e+00  8.065e-01  -2.094 0.036227 *
room_typePrivate room            5.043e-01  4.682e-02  10.770  < 2e-16 ***
room_typeShared room            -6.612e-01  3.622e-01  -1.825 0.067947 .
accommodates                    -1.285e-01  1.494e-02  -8.604  < 2e-16 ***
bathrooms_text                   5.092e-02  1.883e-02   2.705 0.006839 **
amenities                        2.247e-02  1.140e-03  19.703  < 2e-16 ***
price                            2.704e-01  3.681e-02   7.345 2.06e-13 ***
minimum_nights                   9.311e-02  1.284e-02   7.253 4.08e-13 ***
maximum_nights                   1.191e-04  3.254e-05   3.659 0.000253 ***
number_of_reviews_ltm            4.616e-02  1.147e-03  40.256  < 2e-16 ***
instant_bookable1               -1.604e-01  4.011e-02  -3.999 6.37e-05 ***
review_scores                    1.661e+00  3.510e-02  47.320  < 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: 41164  on 45163  degrees of freedom
Residual deviance: 26926  on 45141  degrees of freedom
AIC: 26972

Number of Fisher Scoring iterations: 7
```

<그림4>

```
> summary(mod3) # AIC: 26984

Call:
glm(formula = host_is_superhost ~ description + host_since +
    host_about + host_response_time + host_response_rate + host_total_listings_count +
    host_has_profile_pic + room_type + accommodates + bathrooms_text +
    amenities + price + minimum_nights + number_of_reviews_ltm +
    instant_bookable + review_scores, family = binomial(link = "logit"),
    data = london)

Coefficients:
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                         -4.840e+01  9.264e-01 -52.251  < 2e-16 ***
description                          3.348e-03  5.401e-04   6.199 5.67e-10 ***
host_since                           1.254e-04  1.404e-05   8.936  < 2e-16 ***
host_about                           4.890e-03  5.498e-04   8.895  < 2e-16 ***
host_response_timeNoInfo             6.883e-01  2.773e-01   2.482  0.01305 *
host_response_timewithin a day      -7.408e-01  3.384e-01  -2.189  0.02859 *
host_response_timewithin a few hours -4.480e-01  3.455e-01  -1.297  0.19474
host_response_timewithin an hour    -1.596e-01  3.467e-01  -0.460  0.64539
host_response_rate                   3.151e-02  3.043e-03  10.356  < 2e-16 ***
host_total_listings_count           -5.837e-03  3.946e-03  -1.479  0.13907
host_has_profile_pic1                2.954e-01  1.159e-01   2.548  0.01083 *
room_typeHotel room                 -1.650e+00  8.053e-01  -2.049  0.04042 *
room_typePrivate room                5.023e-01  4.681e-02  10.730  < 2e-16 ***
room_typeShared room                -6.519e-01  3.622e-01  -1.800  0.07192 .
accommodates                        -1.283e-01  1.493e-02  -8.594  < 2e-16 ***
bathrooms_text                       5.205e-02  1.882e-02   2.766  0.00567 **
amenities                            2.253e-02  1.139e-03  19.768  < 2e-16 ***
price                                2.708e-01  3.681e-02   7.355 1.91e-13 ***
minimum_nights                       9.238e-02  1.283e-02   7.198 6.13e-13 ***
number_of_reviews_ltm                4.618e-02  1.146e-03  40.307  < 2e-16 ***
instant_bookable1                   -1.599e-01  4.010e-02  -3.987 6.70e-05 ***
review_scores                        1.651e+00  3.493e-02  47.281  < 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: 41164  on 45163  degrees of freedom
Residual deviance: 26940  on 45142  degrees of freedom
AIC: 26984

Number of Fisher Scoring iterations: 7
```

<그림5>