

EXECUTIVE SUMMARY

Property that has shared spaces, and share bathrooms have lower listing prices. Hosts with more experience tend to charge a higher price. The population of the city also negatively impacts the listing price due to fierce competition. White hosts do not have any bias on pricing.

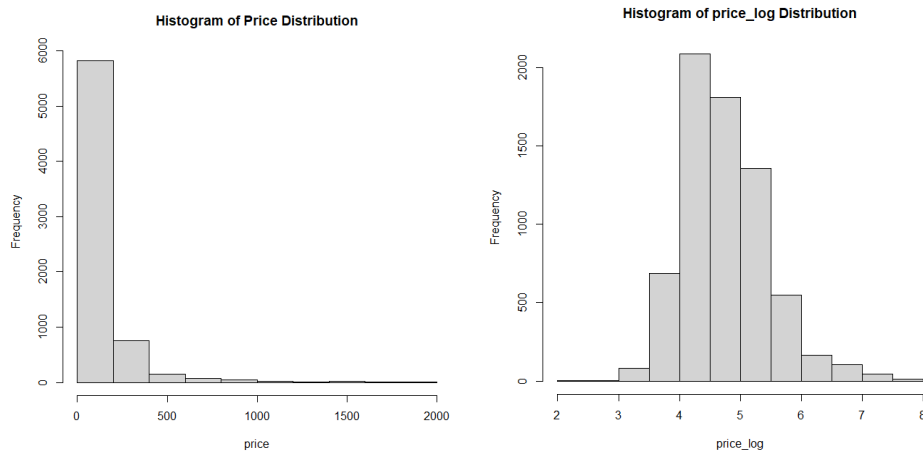
Regarding the demand for properties, there is a higher demand for hosts that make multiple listings, have long experience, are of high age, are male, and have properties located in big cities. There is a lower demand for properties that are not rated, have share areas/bathrooms, include cleaning fees, and are owned by black hosts. There tends to have less demand for 5-star rating properties, which could be explained by their higher price.

Host acceptance rates are also impacted by guest gender and race. White female guests have a 44% higher chance to be accepted by hosts, while that of black females is -35%. Hosts likely to say yes to white male guests (9% higher chance) than black male guests, who are 18% less likely to be accepted.

ANALYSES

Goal 1: Examine what factors affect Airbnb listing prices

Outcome variable selection



- The variable 'price' follows power law distribution which is not suitable for running regression.
- The variable 'log_price' follows a normal distribution, we use this variable to predict the outcome variable.

Initial regression model

Coefficients:					Pr(> t)	
	Estimate	Std. Error	t value			
(Intercept)	4.683615	0.107709	43.484	< 2e-16	***	
shared_property	-0.719021	0.014302	-50.274	< 2e-16	***	
shared_bathroom	-0.188262	0.022447	-8.387	< 2e-16	***	
strict_cancellation	0.121536	0.014621	8.313	< 2e-16	***	
has_cleaning_fee	-0.034968	0.014247	-2.454	0.01414	*	
multiple_listings	-0.010519	0.013683	-0.769	0.44204		
not_rated	0.316393	0.016978	18.636	< 2e-16	***	
five_star_property	0.113175	0.016837	6.722	1.94e-11	***	
log_host_experience	0.048198	0.006581	7.324	2.69e-13	***	
log_tract_listings	0.088142	0.006441	13.684	< 2e-16	***	
log_population	-0.029606	0.011643	-2.543	0.01102	*	
host_age_old	0.007376	0.031414	0.235	0.81437		
host_age_unknown	0.017665	0.039361	0.449	0.65361		
host_age_young	-0.067408	0.013727	-4.911	9.28e-07	***	
host_gender_FF	-0.064485	0.044478	-1.450	0.14716		
host_gender_M	0.038746	0.015179	2.553	0.01071	*	
host_gender_MF	0.017208	0.019518	0.882	0.37798		
host_gender_MM	0.038512	0.045380	0.849	0.39610		
host_gender_unknown	0.100088	0.035054	2.855	0.00431	**	
host_race_black	-0.033375	0.033921	-0.984	0.32520		
host_race_hisp	0.005358	0.044987	0.119	0.90519		
host_race_multiracial	-0.030291	0.047194	-0.642	0.52099		
host_race_unknown	0.009308	0.035324	0.263	0.79218		
host_race_white	0.053961	0.028224	1.912	0.05593	.	
city_Baltimore	0.227912	0.034099	6.684	2.51e-11	***	
city_Dallas	0.076637	0.032437	2.363	0.01817	*	
city_Los-Angeles	0.227451	0.021865	10.403	< 2e-16	***	
city_St-Louis	-0.033826	0.046580	-0.726	0.46774		
city_Washington	0.157455	0.023543	6.688	2.44e-11	***	

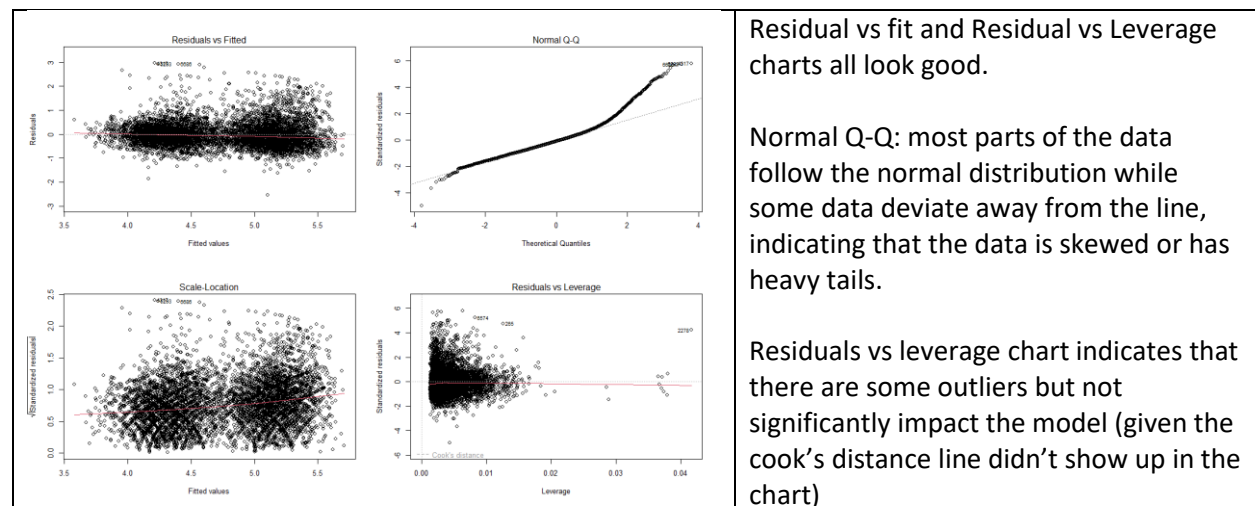
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						
Residual standard error: 0.5119 on 6864 degrees of freedom						
Multiple R-squared: 0.4496, Adjusted R-squared: 0.4474						
F-statistic: 200.3 on 28 and 6864 DF, p-value: < 2.2e-16						

Observation on variables:

- 'Shared_property' has a negative impact on the price of the listing house.
- 'Log_host_experience' implies that hosts with more experience charge higher prices.
- 'Log_population' negatively impacts price, which could be due to competition in the big city

The Adjusted R-squared indicates that the model explain 0.4474 of the 'log_price'

Model diagnostics



Residual vs fit and Residual vs Leverage charts all look good.

Normal Q-Q: most parts of the data follow the normal distribution while some data deviate away from the line, indicating that the data is skewed or has heavy tails.

Residuals vs leverage chart indicates that there are some outliers but not significantly impact the model (given the cook's distance line didn't show up in the chart)

Multicollinearity

<div><h3>VIF stepwise with the threshold of 5</h3><p>There is 1 variable that has a collinearity problem with the host_race_white</p></div>	<div><h3>Lasso model with lambda = 10</h3><p>Beta of all variables equal zero which indicate that all variables are eligible for the model</p></div>																																																																																																																																					
<div><h4>Output:</h4><p>1 variables from the 25 input variables have collinearity problem:</p><p>host_race_white</p><p>After excluding the collinear variables, the linear correlation coefficients ranges between:</p><p>min correlation (city_Los-Angeles ~ host_gender_M): -3.122493e-05</p><p>max correlation (host_race_unknown ~ host_gender_unknown): 0.6620784</p><p>----- VIFs of the remained variables -----</p><table><thead><tr><th></th><th>Variables</th><th>VIF</th></tr></thead><tbody><tr><td>1</td><td>log_price</td><td>1.742021</td></tr><tr><td>2</td><td>shared_property</td><td>1.820345</td></tr><tr><td>3</td><td>shared_bathroom</td><td>1.153677</td></tr><tr><td>4</td><td>strict_cancellation</td><td>1.185333</td></tr><tr><td>5</td><td>has_cleaning_fee</td><td>1.235558</td></tr><tr><td>6</td><td>multiple_listings</td><td>1.079347</td></tr><tr><td>7</td><td>not_rated</td><td>1.838286</td></tr><tr><td>8</td><td>five_star_property</td><td>1.728289</td></tr><tr><td>9</td><td>host_age_old</td><td>1.054747</td></tr><tr><td>10</td><td>host_age_young</td><td>1.088879</td></tr><tr><td>11</td><td>host_gender_FF</td><td>1.044645</td></tr><tr><td>12</td><td>host_gender_M</td><td>1.276893</td></tr><tr><td>13</td><td>host_gender_MF</td><td>1.277343</td></tr><tr><td>14</td><td>host_gender_MM</td><td>1.067747</td></tr><tr><td>15</td><td>host_gender_unknown</td><td>2.049287</td></tr><tr><td>16</td><td>host_race_black</td><td>1.118437</td></tr><tr><td>17</td><td>host_race_hisp</td><td>1.022057</td></tr><tr><td>18</td><td>host_race_multiracial</td><td>1.090844</td></tr><tr><td>19</td><td>host_race_unknown</td><td>1.856043</td></tr><tr><td>20</td><td>city_Baltimore</td><td>1.418541</td></tr><tr><td>21</td><td>city_Dallas</td><td>1.471399</td></tr><tr><td>22</td><td>city_Los-Angeles</td><td>3.054821</td></tr><tr><td>23</td><td>city_St-Louis</td><td>1.188300</td></tr><tr><td>24</td><td>city_Washington</td><td>2.447989</td></tr></tbody></table></div>		Variables	VIF	1	log_price	1.742021	2	shared_property	1.820345	3	shared_bathroom	1.153677	4	strict_cancellation	1.185333	5	has_cleaning_fee	1.235558	6	multiple_listings	1.079347	7	not_rated	1.838286	8	five_star_property	1.728289	9	host_age_old	1.054747	10	host_age_young	1.088879	11	host_gender_FF	1.044645	12	host_gender_M	1.276893	13	host_gender_MF	1.277343	14	host_gender_MM	1.067747	15	host_gender_unknown	2.049287	16	host_race_black	1.118437	17	host_race_hisp	1.022057	18	host_race_multiracial	1.090844	19	host_race_unknown	1.856043	20	city_Baltimore	1.418541	21	city_Dallas	1.471399	22	city_Los-Angeles	3.054821	23	city_St-Louis	1.188300	24	city_Washington	2.447989	<div><h4>Output:</h4><p>28 x 1 sparse Matrix of class "dgCMatrix"</p><table><thead><tr><th></th><th>s0</th></tr></thead><tbody><tr><td>shared_property</td><td>0</td></tr><tr><td>shared_bathroom</td><td>.</td></tr><tr><td>strict_cancellation</td><td>.</td></tr><tr><td>has_cleaning_fee</td><td>.</td></tr><tr><td>multiple_listings</td><td>.</td></tr><tr><td>not_rated</td><td>.</td></tr><tr><td>five_star_property</td><td>.</td></tr><tr><td>log_host_experience</td><td>.</td></tr><tr><td>log_tract_listings</td><td>.</td></tr><tr><td>log_population</td><td>.</td></tr><tr><td>host_age_old</td><td>.</td></tr><tr><td>host_age_unknown</td><td>.</td></tr><tr><td>host_age_young</td><td>.</td></tr><tr><td>host_gender_FF</td><td>.</td></tr><tr><td>host_gender_M</td><td>.</td></tr><tr><td>host_gender_MF</td><td>.</td></tr><tr><td>host_gender_MM</td><td>.</td></tr><tr><td>host_gender_unknown</td><td>.</td></tr><tr><td>host_race_black</td><td>.</td></tr><tr><td>host_race_hisp</td><td>.</td></tr><tr><td>host_race_multiracial</td><td>.</td></tr><tr><td>host_race_unknown</td><td>.</td></tr><tr><td>host_race_white</td><td>.</td></tr><tr><td>city_Baltimore</td><td>.</td></tr><tr><td>city_Dallas</td><td>.</td></tr><tr><td>city_Los-Angeles</td><td>.</td></tr><tr><td>city_St-Louis</td><td>.</td></tr><tr><td>city_Washington</td><td>.</td></tr></tbody></table></div>		s0	shared_property	0	shared_bathroom	.	strict_cancellation	.	has_cleaning_fee	.	multiple_listings	.	not_rated	.	five_star_property	.	log_host_experience	.	log_tract_listings	.	log_population	.	host_age_old	.	host_age_unknown	.	host_age_young	.	host_gender_FF	.	host_gender_M	.	host_gender_MF	.	host_gender_MM	.	host_gender_unknown	.	host_race_black	.	host_race_hisp	.	host_race_multiracial	.	host_race_unknown	.	host_race_white	.	city_Baltimore	.	city_Dallas	.	city_Los-Angeles	.	city_St-Louis	.	city_Washington	.
	Variables	VIF																																																																																																																																				
1	log_price	1.742021																																																																																																																																				
2	shared_property	1.820345																																																																																																																																				
3	shared_bathroom	1.153677																																																																																																																																				
4	strict_cancellation	1.185333																																																																																																																																				
5	has_cleaning_fee	1.235558																																																																																																																																				
6	multiple_listings	1.079347																																																																																																																																				
7	not_rated	1.838286																																																																																																																																				
8	five_star_property	1.728289																																																																																																																																				
9	host_age_old	1.054747																																																																																																																																				
10	host_age_young	1.088879																																																																																																																																				
11	host_gender_FF	1.044645																																																																																																																																				
12	host_gender_M	1.276893																																																																																																																																				
13	host_gender_MF	1.277343																																																																																																																																				
14	host_gender_MM	1.067747																																																																																																																																				
15	host_gender_unknown	2.049287																																																																																																																																				
16	host_race_black	1.118437																																																																																																																																				
17	host_race_hisp	1.022057																																																																																																																																				
18	host_race_multiracial	1.090844																																																																																																																																				
19	host_race_unknown	1.856043																																																																																																																																				
20	city_Baltimore	1.418541																																																																																																																																				
21	city_Dallas	1.471399																																																																																																																																				
22	city_Los-Angeles	3.054821																																																																																																																																				
23	city_St-Louis	1.188300																																																																																																																																				
24	city_Washington	2.447989																																																																																																																																				
	s0																																																																																																																																					
shared_property	0																																																																																																																																					
shared_bathroom	.																																																																																																																																					
strict_cancellation	.																																																																																																																																					
has_cleaning_fee	.																																																																																																																																					
multiple_listings	.																																																																																																																																					
not_rated	.																																																																																																																																					
five_star_property	.																																																																																																																																					
log_host_experience	.																																																																																																																																					
log_tract_listings	.																																																																																																																																					
log_population	.																																																																																																																																					
host_age_old	.																																																																																																																																					
host_age_unknown	.																																																																																																																																					
host_age_young	.																																																																																																																																					
host_gender_FF	.																																																																																																																																					
host_gender_M	.																																																																																																																																					
host_gender_MF	.																																																																																																																																					
host_gender_MM	.																																																																																																																																					
host_gender_unknown	.																																																																																																																																					
host_race_black	.																																																																																																																																					
host_race_hisp	.																																																																																																																																					
host_race_multiracial	.																																																																																																																																					
host_race_unknown	.																																																																																																																																					
host_race_white	.																																																																																																																																					
city_Baltimore	.																																																																																																																																					
city_Dallas	.																																																																																																																																					
city_Los-Angeles	.																																																																																																																																					
city_St-Louis	.																																																																																																																																					
city_Washington	.																																																																																																																																					

Regression model without “host_race_white” variable model

Coefficients:					Observation
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	4.683615	0.107709	43.484	< 2e-16 ***	The adjusted R-squared stay the same, indicates that the “host_race_white” does not help improve model performance
shared_property	-0.719021	0.014302	-50.274	< 2e-16 ***	
shared_bathroom	-0.188262	0.022447	-8.387	< 2e-16 ***	
strict_cancellation	0.121536	0.014621	8.313	< 2e-16 ***	
has_cleaning_fee	-0.034968	0.014247	-2.454	0.01414 *	
multiple_listings	-0.010519	0.013683	-0.769	0.44204	
not_rated	0.316393	0.016978	18.636	< 2e-16 ***	
five_star_property	0.113175	0.016837	6.722	1.94e-11 ***	
log_host_experience	0.048198	0.006581	7.324	2.69e-13 ***	
log_tract_listings	0.088142	0.006441	13.684	< 2e-16 ***	
log_population	-0.029606	0.011643	-2.543	0.01102 *	
host_age_old	0.007376	0.031414	0.235	0.81437	
host_age_unknown	0.017665	0.039361	0.449	0.65361	
host_age_young	-0.067408	0.013727	-4.911	9.28e-07 ***	
host_gender_FF	-0.064485	0.044478	-1.450	0.14716	
host_gender_M	0.038746	0.015179	2.553	0.01071 *	
host_gender_MF	0.017208	0.019518	0.882	0.37798	
host_gender_MM	0.038512	0.045380	0.849	0.39610	
host_gender_unknown	0.100088	0.035054	2.855	0.00431 **	
host_race_black	-0.033375	0.033921	-0.984	0.32520	
host_race_hisp	0.005358	0.044987	0.119	0.90519	
host_race_multiracial	-0.030291	0.047194	-0.642	0.52099	
host_race_unknown	0.009308	0.035324	0.263	0.79218	
host_race_white	0.053961	0.028224	1.912	0.05593 .	
city_Baltimore	0.227912	0.034099	6.684	2.51e-11 ***	
city_Dallas	0.076637	0.032437	2.363	0.01817 *	
city_Los-Angeles	0.227451	0.021865	10.403	< 2e-16 ***	
city_St-Louis	-0.033826	0.046580	-0.726	0.46774	
city_Washington	0.157455	0.023543	6.688	2.44e-11 ***	
--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Residual standard error: 0.5119 on 6864 degrees of freedom Multiple R-squared: 0.4496, Adjusted R-squared: 0.4474 F-statistic: 200.3 on 28 and 6864 DF, p-value: < 2.2e-16					

Goal 2: Analyze how prices and other characteristics impact the total number of guests.

Poisson regression “total_guests” ~ “log_price”

Result					Interpretation	
Coefficients:					(Intercept)	log_price
(Intercept)	3.709970	0.014280	259.80	<2e-16 ***	40.8525751	0.9667636
log_price	-0.033801	0.002994	-11.29	<2e-16 ***		
Observation:						
Price negatively impacts demand. One unit increase in log_price will reduce demand by 3.4%						

Poisson regression “total_guests” ~ all other characteristics variables

Multivariate Poisson regression of “log_total_guest”					Interpretation			
Coefficients:					(Intercept)	log_price	shared_property	shared_bathroom
(Intercept)	3.810481	0.041486	91.849	< 2e-16 ***	45.1721627	0.6536443	0.6401305	0.7262517
log_price	-0.425192	0.004740	-89.695	< 2e-16 ***	strict_cancellation	has_cleaning_fee	multiple_listings	not_rated
shared_property	-0.446083	0.005754	-77.532	< 2e-16 ***	1.1630416	0.8144122	3.2548604	0.3064875
shared_bathroom	-0.319859	0.009129	-35.038	< 2e-16 ***	five_star_property	log_host_experience	log_tract_listings	log_population
strict_cancellation	0.151039	0.004500	33.566	< 2e-16 ***	0.8904176	1.6811652	1.1670618	0.9419287
has_cleaning_fee	-0.205289	0.005007	-41.004	< 2e-16 ***	host_age_old	host_age_unknown	host_age_young	host_gender_FF
multiple_listings	1.180149	0.004554	259.161	< 2e-16 ***	0.9324638	1.2204082	0.9615404	0.9435309
not_rated	-1.182578	0.006883	-171.810	< 2e-16 ***	host_gender_M	host_gender_MF	host_gender_MM	host_gender_unknown
five_star_property	-0.116065	0.004499	-25.796	< 2e-16 ***	1.0529084	1.1797530	1.5301377	1.0977403
log_host_experience	0.519487	0.002996	173.416	< 2e-16 ***	host_race_black	host_race_hisp	host_race_multiracial	host_race_unknown
log_tract_listings	0.154489	0.002209	69.946	< 2e-16 ***	0.6792353	0.9455598	0.8048147	0.8895910
log_population	-0.059826	0.003617	-16.540	< 2e-16 ***	host_race_white	city_Baltimore	city_Dallas	city_Los-Angeles
host_age_old	-0.069925	0.011495	-6.083	1.18e-09 ***	0.9466718	1.4782298	1.1513473	1.7862672
host_age_unknown	0.199185	0.012556	15.864	< 2e-16 ***	city_St-Louis	city_Washington		
host_age_young	-0.039219	0.004611	-8.506	< 2e-16 ***	1.7839869	1.4583427		
host_gender_FF	-0.058126	0.015053	-3.861	0.000113 ***				
host_gender_M	0.051556	0.005236	9.846	< 2e-16 ***				
host_gender_MF	0.165305	0.006196	26.679	< 2e-16 ***				
host_gender_MM	0.425358	0.013451	31.624	< 2e-16 ***				
host_gender_unknown	0.093254	0.011188	8.335	< 2e-16 ***				
host_race_black	-0.386788	0.013684	-28.265	< 2e-16 ***				
host_race_hisp	-0.055978	0.014703	-3.807	0.000141 ***				
host_race_multiracial	-0.217143	0.015563	-13.953	< 2e-16 ***				
host_race_unknown	-0.116993	0.012228	-9.568	< 2e-16 ***				
host_race_white	-0.054803	0.009527	-5.752	8.80e-09 ***				
city_Baltimore	0.390845	0.013688	28.555	< 2e-16 ***				
city_Dallas	0.140933	0.015539	9.070	< 2e-16 ***				
city_Los-Angeles	0.580128	0.009053	64.084	< 2e-16 ***				
city_St-Louis	0.578851	0.015841	36.541	< 2e-16 ***				
city_Washington	0.377301	0.009819	38.425	< 2e-16 ***				
Observation:					<ul style="list-style-type: none"> Some of the characteristics that greatly improve the demand are multiple listings, long host experience, host age, host gender M/MF/MM, and locations in the big city. Some of the characteristics that reduce the demand are price, not_rate, shared property, shared bathroom, cleaning fee, and host race black. Interestingly, a five-star rating reduces the demand, which could due to the high price of high-rating properties. 			

Compare model fit between one variable and multivariable Poisson regressions

	Pois1	Pois1m	Observation: Multivariable Poisson regression better explain the impact of other variables on demand
AIC	530649.6	259593.7	
BIC	530663.3	259798.8	
logLik	-265322.8	-129766.8	

Goal 3: Examine whether there is discrimination of host response towards gender and race

Probability of host accepting the request by guest gender and race

Results:			Observation:
guest_gender	guest_race	acceptance_prob	
<chr>	<chr>	<dbl>	
1 female	black	0.393	<ul style="list-style-type: none"> Gender: Female customers have higher acceptance rates than male customers. Race: Black customers have lower acceptance rates than white customers Hosts tend to favor white female customers and are less likely to accept the request by black male customers
2 female	white	0.484	
3 male	black	0.346	
4 male	white	0.416	

Logistic regression to explain the outcome variable (host_response_yes) using guest characteristics

Result:

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.41497	0.04244	-9.779	< 2e-16	***
guest_race_white	0.33415	0.04928	6.781	1.19e-11	***
guest_gender_male	-0.24234	0.04926	-4.920	8.66e-07	***

Observation:

- Guest_race_white has higher acceptance change than that of their counter parts.
- Guest_gender_male is less likely to be accepted by the hosts

Logistic regression includes all explanatory variables

Result:

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.02055	0.53149	-3.802	0.000144 ***
shared_property	0.08175	0.06440	1.269	0.204323
shared_bathroom	-0.85116	0.24899	-3.418	0.000630 ***
strict_cancellation	0.15245	0.07343	2.076	0.037885 *
has_cleaning_fee	0.26463	0.09277	2.852	0.004338 **
multiple_listings	0.61196	0.17057	3.588	0.000334 ***
not_rated	-1.18589	0.32279	-3.674	0.000239 ***
five_star_property	-0.24636	0.09443	-2.609	0.009082 **
log_host_experience	-0.30432	0.08576	-3.548	0.000388 ***
log_tract_listings	-0.07931	0.03436	-2.308	0.020984 *
log_population	0.03095	0.04949	0.625	0.531733
host_age_old	0.11157	0.13666	0.816	0.414254
host_age_unknown	-0.07118	0.16677	-0.427	0.669504
host_age_young	-0.20150	0.07850	-2.567	0.010258 *
host_gender_FF	-0.39494	0.21651	-1.824	0.068137 .
host_gender_M	-0.24250	0.09037	-2.683	0.007290 **
host_gender_MF	0.14305	0.09062	1.579	0.114444
host_gender_MM	0.03370	0.18950	0.178	0.858833
host_gender_unknown	0.16031	0.15305	1.047	0.294889
host_race_black	0.71602	0.23700	3.021	0.002518 **
host_race_hisp	-0.09913	0.19046	-0.520	0.602720
host_race_multiracial	0.40332	0.22562	1.788	0.073842 .
host_race_unknown	-0.23958	0.16103	-1.488	0.136795
host_race_white	-0.14532	0.12343	-1.177	0.239074
city_Baltimore	3.85694	0.91106	4.233	2.30e-05 ***
city_Dallas	4.22267	1.00320	4.209	2.56e-05 ***
city_Los-Angeles	3.78176	0.88508	4.273	1.93e-05 ***
city_St-Louis	3.80831	0.90855	4.192	2.77e-05 ***
city_Washington	3.54740	0.82390	4.306	1.67e-05 ***
guest_race_white	0.60471	0.16810	3.597	0.000322 ***
guest_gender_male	-0.36678	0.11047	-3.320	0.000899 ***
pred_prob	-2.09415	1.92002	-1.091	0.275406
predTRUE	-0.12817	0.09325	-1.374	0.169331

The goodness of fit:

	log_0	log_1
AIC	9.268703e+03	8.485592e+03
BIC	9.289218e+03	8.711255e+03
logLik	-4.631351e+03	-4.209796e+03
Pseudo_R2	9.775305e-02	9.775305e-02

Observation:

The regression results confirm that the hosts do accept guests based on gender and race.

The Pseudo R2 value is low, indicating that the model is not robust.

Description of the datasets

(1) Airbnb_listing.csv

- Listing_id: index for listing

Outcome variables

- Price: listing price (prices may vary across dates, consider this price as the representative price or historical average price).
- Log_price: log scaled price
- Total_guests: total number of guests who have booked the property in the past
- Log_total_guests: log scaled total guests

Listing characteristics

- Shared_property: =1 if the property is shared; =0 if the property is private
- Shared_bathroom: =1 if bathroom(s) are shared; =0 if bathroom(s) are private
- Strict_cancellation: =1 if the cancellation rule is strict
- Has_cleaning_fee: =1 if has cleaning fee in addition to the price quoted
- Multiple_listings: =1 if the host has multiple listings on Airbnb
- Not_rated: =1 if the listing has not been rated
- Five_star_property: =1 if the listing has been rated as five stars

Host characteristics

- Host_age: {young, middle, old, unknown}
- Host_gender: {female, male, female-female, male-male, female-male, unknown}
- Host_race: {asian, black, hispanic, white, multiracial, unknown}
- Log_host_experience: log scaled host experience (experience is measured as number of months the host has been on Airnbn)

Location characteristics

- City: listing city
- Log_tract_listing: log scaled number of listings in the corresponding census tract
- Log_population: log scaled total population in the corresponding census tract

(2) Airbnb_request_response.csv

- Listing_id: index for listing

Guest characteristics

- Guest race: {black, white}
- Guest gender: {female, male}

Host response

- Host_response_yes: =1 if the host accepts the request