

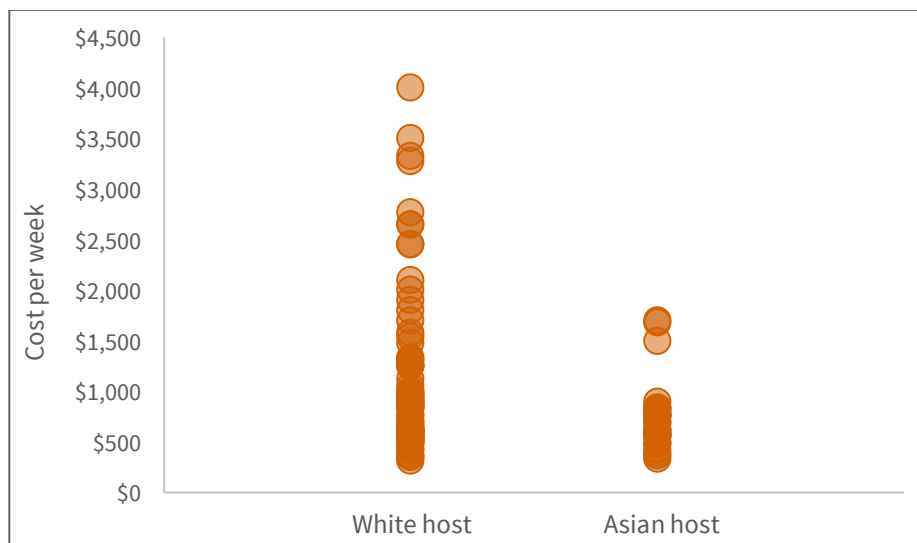


# The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans

David Wang, Stephen Xi, John Gilheany

## Highlights

- We tested if Asians receive lower prices on Airbnb's vacation rental website
- We identified 101 White and Asian hosts on Airbnb in Oakland and Berkeley in April 2015
- We found that on average Asian hosts earn \$90 less per week or 20% less than White hosts for similar rentals



*Price differences between White and Asian hosts on Airbnb*

## Abstract

Benjamin Edelman and Michael Luca of Harvard Business School investigated Airbnb.com, a hospitality exchange service that allows people to rent lodging from hosts who post on the

site [1]. Their findings revealed the prevalence of racial discrimination among hosts — specifically, African American hosts earn much less, on average, than their non-black counterparts in New York. Our paper attempts to build upon this study by rigorously applying statistical techniques to test how Asian American rental incomes compare to those of whites in California.

**Results summary:** Using a scraped dataset from the Oakland/Berkeley affiliate of Airbnb.com, we observed that Asian hosts earn, on average, \$90 (or 20 percent) less per week than white landlords within this location for the standard one-bedroom rental for occupancy of one. The differential increases with the number of bedrooms and other upgrades to the house go up. Our regressions are resilient when subjected to statistical tests. These findings add information to the ongoing discussion on racial biases in commercial transactions.

## Introduction

“On the Internet, nobody knows you are a dog” was a cartoon first published in 1993 [2]. The cartoon suggested that technology and the Internet were indifferent to personal demographics such as race and ethnicity. As technology has become personal and social, however, it is not surprising that societal issues have followed and that two users of different races may not have the same online experience.

Latanya Sweeney was among the first to show an online difference related to race [3]. When searching for results using a person’s name, she found that search engines delivered advertisements that suggested the person had an arrest record more often when the search term was a name primarily associated with black babies compared to searches using names primarily associated with white babies, regardless of whether there were any arrests actually associated with the name.

Benjamin Edelman and Michael Luca of Harvard Business School investigated Airbnb.com, a hospitality exchange service that allows people to rent lodging from hosts who post on the site [4]. On Airbnb, people who have rooms or homes to rent post a description of the facility, its location, price, and pictures of the place and the hosts. Potential renters view the listings to find a place. The research findings of Edelman and Luca revealed the prevalence of racial discrimination in users’ lodging choices. Specifically, African American hosts earn 12 percent less, on average, than their non-black counterparts.

Both of these studies reflected on discrimination against African Americans. Inequality in income between blacks and non-blacks is extensively documented and persists in recent U.S. Census data [5]. Economically, blacks suffer the highest poverty rate and receive the lowest median income of any racial group in America. A National Urban League study revealed that the black-white Equality Index (measuring how non-whites are treated relative to whites) is 55.8 percent in economics and 60.6 percent in social justice [6].

What about online differences experienced by other minority groups? Considered by many as a “model minority”, the Asian American demographic earns the highest median income of all racial groups in America [7], and Asian American students consistently receive the highest standardized test scores [8]. Persons of Asian descent constitute 4.8 percent of the American population, and yet they routinely see strong representation at the nation’s most elite universities and other research institutions [9]. According to the Pew Research Center, Asians recently surpassed Hispanics as the largest immigrant group in the U.S. and are “more satisfied than the general public with their lives, finances and the direction of the country.” [10]

Yet, some also label Asian Americans as a “neglected minority.” On the corporate side, only 30 Fortune 100 companies have Asian American representation on their boards [11]. This has been described as a manifestation of a “bamboo ceiling” phenomenon, whereby Asian Americans are excluded from executive positions based on subjective factors such as “lack of leadership potential and lack of communication skills that cannot actually be explained by job performance or qualifications.” [12] In college admissions, Asians need to score around 140 points higher than whites on the SAT to get into the same top schools and are frequently perceived by admission officers as “featureless and lacking in individuality.” [13] How might this negative image be perpetuated online? While Asian Americans excel statistically on paper, do they face covert racial bias online? How do they do on Airbnb?

## Background

Mentioned earlier, Airbnb is a popular website for people to list, find, and rent lodging [14]. It is a kind of matchmaking service. We asked: How much are consumers willing to pay for an Airbnb rental home if the host is Asian? Does their asking price differ from whites?

Airbnb offers its service internationally, so to answer these questions, we needed to identify a location with parity of Airbnb offerings between the two racial groups. After a careful review, we chose the border area of Oakland/Berkeley, California as the location for analysis because of its diverse demographic profile. While the racial makeup of the United States is heavily skewed towards whites (comprising 77.7 percent of the population [15]), the specific ZIP codes we examined had a particularly even spread of races: 2010 Census figures show that the population is 34.5 percent white, 28 percent black or African American, 25.4 percent Hispanic, and 16.8 percent Asian or Asian American [16].

This area is also relatively socioeconomically diverse. The area’s top employers include the University of California, Berkeley and a number of medical and research facilities — the service industry comprises the majority of employment [17]. The distribution of median household income in this area is relatively uniform, as seen in Figure 1 [18]. While on a regional scale, the Bay Area may be economically diverse, individual neighborhoods may have large differences in socioeconomic levels and thus property prices. We were unable to

regress for ZIP code since it is not a quantifiable figure (it is a categorical variable), and this is one potential limitation of this study.

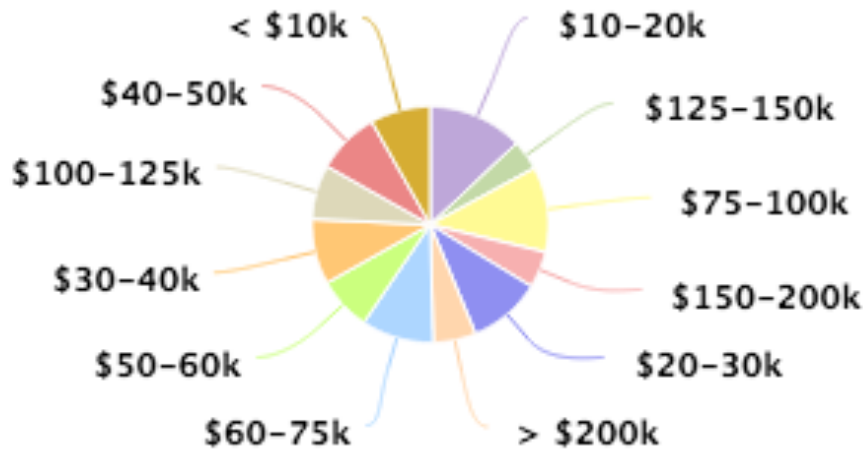


Figure 1: Median income distribution of Oakland/Berkeley area [18]

## Methods

### Basic outline of methods

A myriad of variables can affect the renting price on Airbnb, from number of rooms to number of bathrooms. There is no perfect prediction for housing price because of the noise created by all these factors. But is it possible that race plays a role? Do Asians receive the same amount of money for renting out their homes as do whites, all other things being equal? We aimed to run a series of statistical tests to see if public data from Airbnb.com indicates that Asian race plays a role in rental prices. In doing this, we were able to derive the most appropriate hedonic pricing model with the data set we obtained from Airbnb. Following is a basic outline of the statistical steps we took to achieve a model that predicts how much a host will earn given their race and other factors:

Transformations — use skewness to determine if any transformations are necessary to normalize the data set. Many of our data points required logarithmic and square root transformations.

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.  
<http://techscience.org/a/2015090104/>

Backward Stepwise Regression — informs us which variables have significance in the regression (with an alpha of 0.05).

Diagnosis — tests for skewness and kurtosis (skewness and kurtosis test for normality) and heteroskedasticity (Breusch-Pagan/Cook-Weisberg Test).

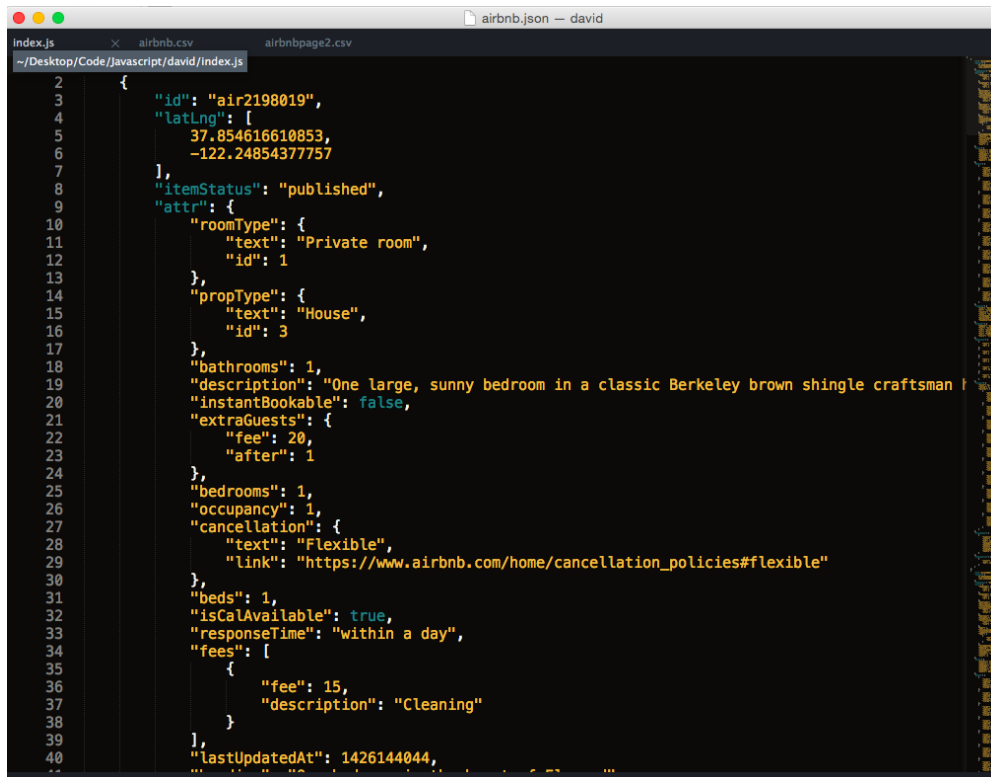
## Obtaining data

To optimize our data results and reduce the risk of human error, we relied on a Mashape API to scrape all necessary information from Oakland/Berkeley Airbnb pages. A small excerpt of the Javascript code for the scraper can be found in Figure 2a. After running the program in the terminal window, the API saves the JSON result to a separate file containing our dataset (Figure 2b). The data set contains key pieces of data including links to the profile picture of the host, number of bathrooms, number of bedrooms, occupancy, and price. Although Airbnb.com lists the daily cost of renting a lodging, we choose to use weekly price because daily price is much too volatile a variable. We gathered our data on April 23<sup>rd</sup> 2015, which included all weeklong stays listed after that date. Since hosts also do not list their race on their Airbnb profiles, we manually sorted through their profile pictures to account for race. We created a dummy variable for race, assigning 0 to Caucasian and 1 to Asian after evaluation of the profile pictures. (Figure 3). We omitted any hosts that did not appear to be white or Asian, or if it was ambiguous what race the host was. Figure 3 includes a profile included in our study, where the photo is a prominent feature of the page. We removed any ambiguous profiles in which we could not clearly discern the host's race.

```
index.js  x  airbnb.csv
1  var cheerio = require('cheerio');
2  var request = require('request-ssl');
3  var fs = require('fs');
4  var _ = require('underscore');
5  var unirest = require('unirest')
6
7
8  // These code snippets use an open-source library. http://unirest.io/nodejs
9  unirest.get("https://zilyo.p.mashape.com/search?latitude=37.854&longitude=-122.256page=2&resultsperpag
10 .header("X-Mashape-Key", "QBRlu6Ic5VmshhwZeA23dKgdmWsnp1IRtQajsnxDLgpkVGDLw0")
11 .header("Accept", "application/json")
12 .end(function (result) {
13   console.log(JSON.parse(result.body).result.length);
14
15   fs.writeFile('airbnbpage2.json',
16     JSON.stringify(JSON.parse(result.body).result, null, 4));
17 });
```

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.  
<http://techscience.org/a/2015090104/>

Figure 2a: Javascript data scraper



```
2  {
3    "id": "air2198019",
4    "latLng": [
5      37.854616610853,
6      -122.24854377757
7    ],
8    "itemStatus": "published",
9    "attr": {
10     "roomType": {
11       "text": "Private room",
12       "id": 1
13     },
14     "propType": {
15       "text": "House",
16       "id": 3
17     },
18     "bathrooms": 1,
19     "description": "One large, sunny bedroom in a classic Berkeley brown shingle craftsman",
20     "instantBookable": false,
21     "extraGuests": {
22       "fee": 20,
23       "after": 1
24     },
25     "bedrooms": 1,
26     "occupancy": 1,
27     "cancellation": {
28       "text": "Flexible",
29       "link": "https://www.airbnb.com/home/cancellation_policies#flexible"
30     },
31     "beds": 1,
32     "isCalAvailable": true,
33     "responseTime": "within a day",
34     "fees": [
35       {
36         "fee": 15,
37         "description": "Cleaning"
38       }
39     ],
40     "lastUpdatedAt": 1426144044,
```

Figure 2b: Scraped data set

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.  
<http://techscience.org/a/2015090104/>

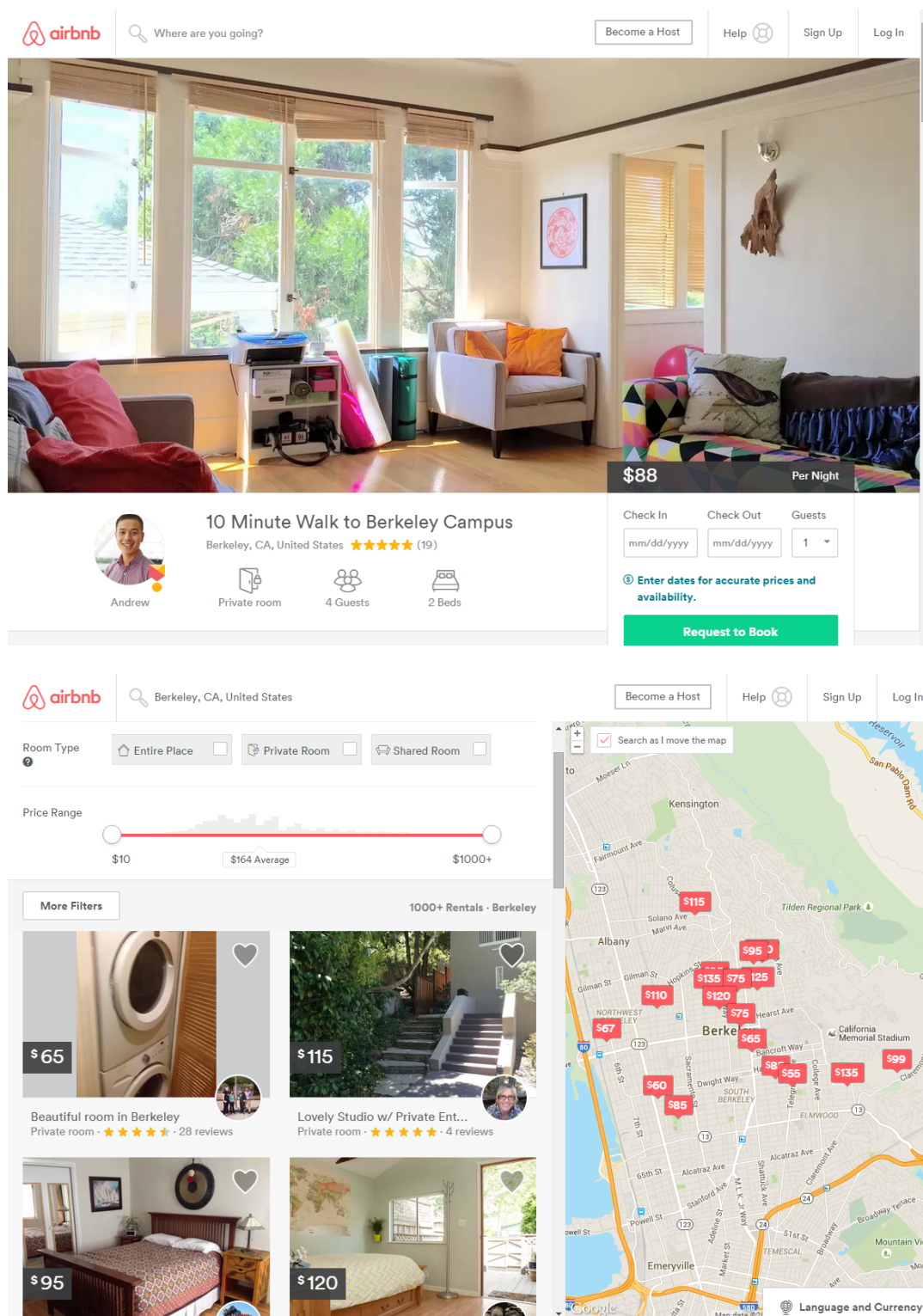


Figure 3: The display page of a scraped URL, depicting the “profile picture”



In order to load our data into Stata, we converted the data set from its coded JavaScript form into an Excel spreadsheet (Figure 4). Our final data set had an overall sample size of 101, with 75 hosts who appeared to be white and 26 hosts who appeared to be of Asian descent.

	A	B	C	D	E
1	bathrooms	bedrooms	occupancy	price	race
2	1	1	2	965	0
3	1	1	4	1008	0
4	1	1	2	650	0
5	1	1	2	567	1
6	2	3	8	840	1
7	1	2	3	825	1
8	2.5	3	8	2100	0
9	1	1	2	850	0
10	2	3	8	2000	0
11	3.5	5	8	3500	0
12	1	1	4	950	0
13	1	1	1	500	0
14	1	2	4	580	1
15	1	1	2	850	0
16	1	0	2	600	0
17	1	3	6	2765	0
18	1	1	4	910	0
19	1	1	1	525	0
20	1	2	4	1120	0
21	1	1	2	615	1

**Figure 4: Excerpt from converted excel sheet of data set**

### Summarizing data

With this data set in hand, we imported the Excel file into Stata, selecting the first row as variable names. Stata provided us with a statistical summary of this data (see Figure 5). We initialized the price as the dependent variable and race, number of bedrooms, number of bathrooms, and occupancy as the independent variables. With a careful selection of the location box while scraping the data in addition to the homogeneity of median incomes, we proceeded with the assumption that the location will not significantly impact the final outcome. However, before we ran the stepwise regression we wanted to observe the relationship between all variables. The most powerful tool to visualize this is a scatterplot matrix (Figure 6).



**. summarize**

Variable	Obs	Mean	Std. Dev.	Min	Max
bathrooms	101	1.282178	.6797641	0	3.5
bedrooms	101	1.534653	1.091461	0	5
occupancy	101	3.584158	2.307671	1	12
price	101	1032.436	750.3322	315	4000
race	101	.2574257	.4393965	0	1
lprice	101	6.749554	.5841823	5.752573	8.294049
lbedrooms	93	.3650973	.5085237	0	1.609438
sqrtrlbedro~s	93	.3663369	.4831193	0	1.268636
bathroomssq	101	2.101485	2.274912	1	12.25
lbathroomssq	101	.4093993	.7172012	0	2.505526
loccupancy	101	1.088224	.6165658	0	2.484907
res	93	4.43e-10	.3348048	-.7423721	.6469355

Figure 5: Statistical summary of data set

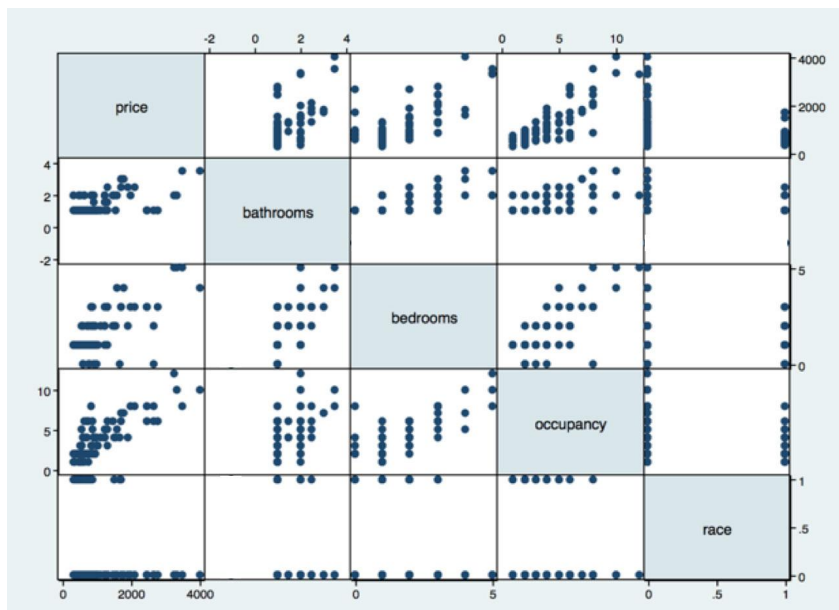


Figure 6: Scatterplot visualization

Figure 6 displays all the scatterplots of the entire dataset, showing the relation of each variable to each other variable. For our purposes, we looked at sale price and its graphical relation to each x-variable (bedrooms, occupancy, etc.). These scatterplots serve as a visual representation to give a general idea of trends and outliers. From the plots, it is clear that some of the relations are clustered or non-linear (such as a “U” shape seen in price vs. occupancy) and therefore require transformation in addition to removal of outliers from the dataset. From the scatterplots, we noted that some outliers included points that were either fairly isolated or influential.

Although the scatterplot matrix gives a good general idea of which variables are skewed, we needed to obtain quantitative evidence of skewness. To concretely see the distribution of the variables and how skewed each one is, we proceeded to run the kernel density estimate plotted against a normal (Figure 7). We categorized the type of skewness of each variable, for example right-skewed or left-skewed, for the transformation to make them more normal. Although each variable is tailored to its own properties, a common trend in our data is a strong right-skewed distribution.

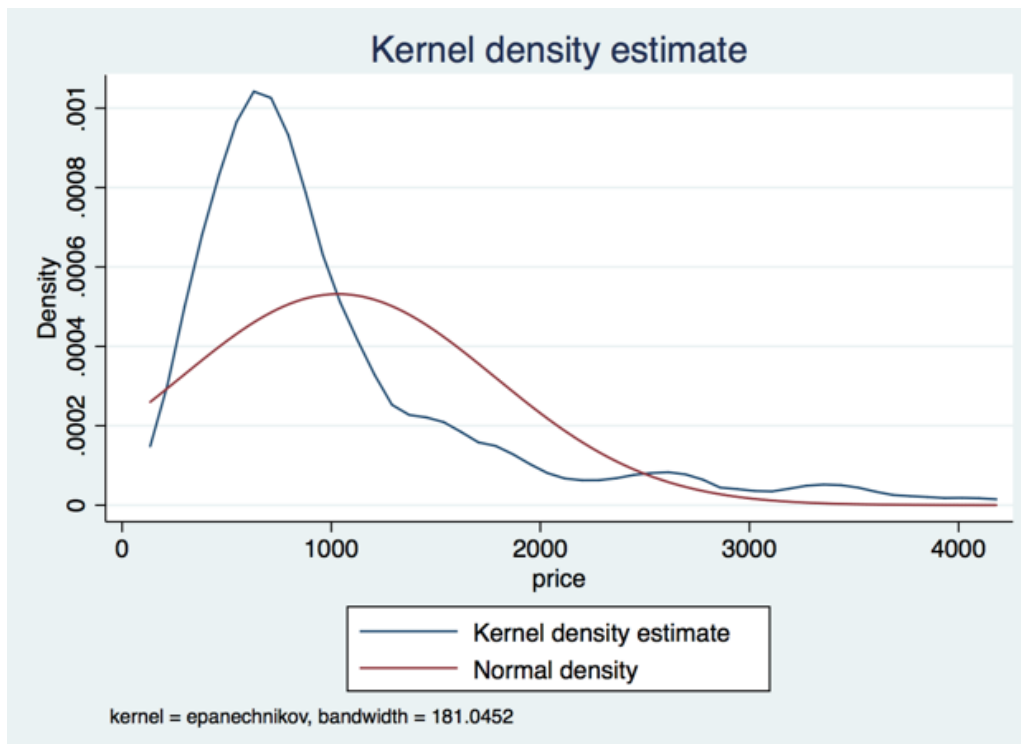
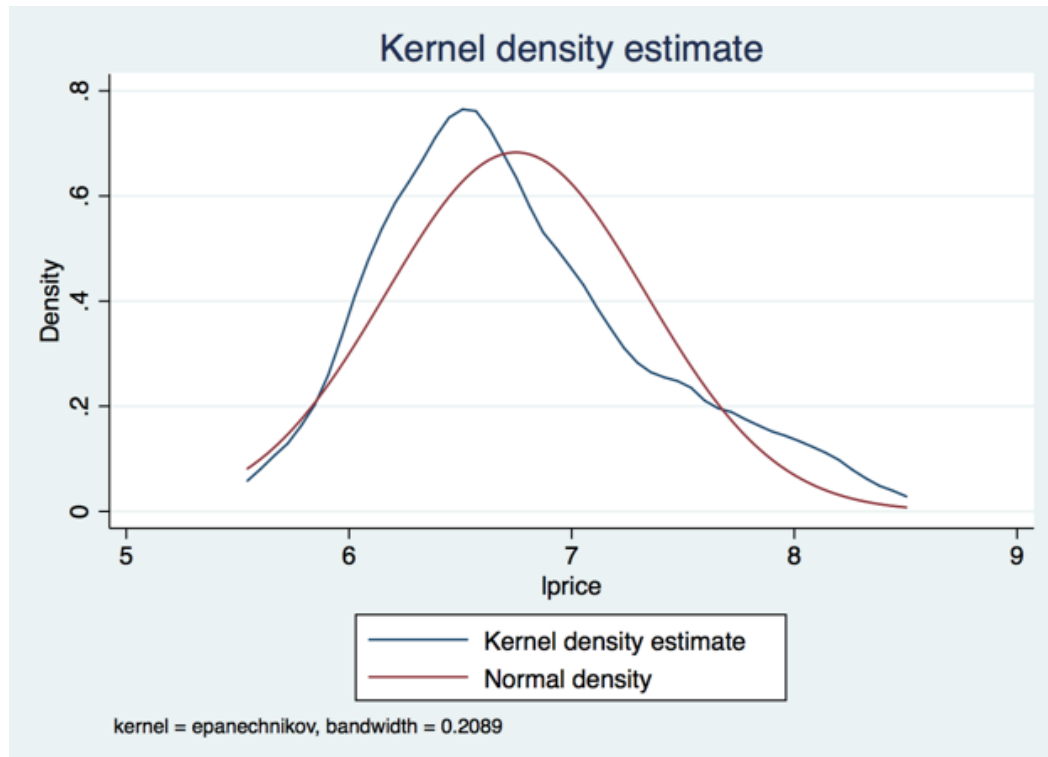


Figure 7: Kernel density estimate of (price) plotted against the normal

Backward stepwise regression and transformations

To determine what types of transformations were necessary, we visualized our data points through the kernel density estimate to quantify skewness and other properties of the data. There were a multitude of transformations we could have used, including cube, square, identity, square root, cube root, and logarithmic. Many of our data points required logarithmic or square root transformations (Figure 8).



**Figure 8: Kernel density estimate of  $\ln(\text{price})$  plotted against the normal**

After trial and error, in addition to statistical intuition, we transformed the data into a more normalized set (Figure 9). The result was a model that was the most optimal set of data transformations (Figure 10). Because the p-value for bathrooms is greater than 0.05, we rejected the null hypothesis that this is a significant variable and omitted it from subsequent statistical processes.

```
. generate lprice = log(price)

. generate lbedrooms = log(bedrooms)
(8 missing values generated)

. generate sqrtlbedrooms = sqrt(lbedrooms)
(8 missing values generated)

. generate bathroomssq = bathrooms*bathrooms

. generate lbathroomssq = log(bathroomssq)

. generate loccupancy = log(occupancy)

. generate racesq = race*race
```

Figure 9: Transformation of variables to normalize the data

```
. sw regress lprice sqrtlbedrooms lbathroomssq loccupancy racesq, pr(0.05)
              begin with full model
p = 0.3353 >= 0.0500 removing lbathroomssq
```

Source	SS	df	MS	Number of obs =	93
Model	21.7860666	3	7.2620222	F( 3, 89) =	62.67
Residual	10.3126741	89	.115872743	Prob > F =	0.0000
Total	32.0987407	92	.348899356	R-squared =	0.6787
				Adj R-squared =	0.6679
				Root MSE =	.3404

lprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
sqrtlbedrooms	.5724729	.1187055	4.82	0.000	.3366076	.8083382
racesq	-.2013149	.0835426	-2.41	0.018	-.3673123	-.0353176
loccupancy	.3438338	.0926322	3.71	0.000	.1597757	.5278919
_cons	6.20077	.086906	71.35	0.000	6.02809	6.37345

Figure 10: Stepwise regression through a sequence of t-tests with significance level 0.05

## Results

### The Data Set

The dataset consisted of 101 hosts, 75 labeled as white and 26 as Asian. There were 74 1 bathrooms (52 from white hosts and 22 from Asian hosts). The average number of bathrooms is 1.32 (1.37 for white hosts and 1.17 for Asian hosts). The average number of bedrooms is 1.57 (1.68 for white hosts and 1.27 for Asian hosts). The smallest rooming option had 0 bedrooms and 1 baths by a white host (for example couch in living room or futon in family room). The largest rooming option had had 5 bedrooms and 3.5 baths by a white host. Figure 5 has a summary.

### Hedonic pricing model

From the above processes we were finally able to derive an equation for determining the sale price of a given host (Figure 11). This serves as a robust representation for finding rental price of a lodging based on bedrooms, race, and occupancy. After careful mathematical manipulation and calculation, we reached the following conclusion: Assuming the bare-minimum setup of a one-bedroom rental for occupancy of one person, Asian Americans, on average, earn \$89.72 less per week than their white counterparts. If the rental were for two bedrooms for one person, for example, our model predicts that the difference would be even greater, with Asian Americans making on average \$144.45 (or 20 percent) less per week than Caucasians.

$$\ln(\text{price}) = 0.572 * \sqrt{\ln(\text{bedrooms})} + 0.344 * \ln(\text{occupancy}) - 0.201 * \text{race}^2 + 6.200$$

$$\text{price} = e^{0.572 * \sqrt{\ln(\text{bedrooms})} + 0.344 * \ln(\text{occupancy}) - 0.201 * \text{race}^2 + 6.200}$$

**Figure 11: Equation for estimating price differential**

After reaching this conclusion, we wanted to run a couple of diagnostic tests to make sure our data did not require additional transformations and that our error was normal.

### Diagnostic tests

Our first test involved testing skewness and kurtosis for normality. By utilizing Stata to calculate residuals (deviation of the observed value from the actual value), we can test for the normality of error. The null hypothesis in this case is that the noise is normal. After running this test, we derived a p-value of 0.1214 (Figure 12). Because this is greater than 0.05, we failed to reject the null hypothesis and assume there is no evidence of non-normality, allowing the test to pass.

```
. predict res, r
(8 missing values generated)

. sktest res
```

Skewness/Kurtosis tests for Normality					
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
res	93	0.5043	0.0562	4.22	0.1214

**Figure 12: Diagnostic Test #1 — Skewness/Kurtosis test for normality**

Our second test assesses heteroskedasticity, which is essentially variability. The null hypothesis here states that there is constant variance (what we were striving for). Our resulting p-value of 0.1272 is much larger than 0.05, so we failed to reject the null hypothesis — thus, we could assume the model does not have heteroskedastic noise (Figure 13). This test also passed for our model.

```
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of lprice

chi2(1)      =    2.33
Prob > chi2   =    0.1272
```

**Figure 13: Diagnostic Test #2 — Cook-Weisburg test for heteroskedasticity**

## Takeaways

With the exclusion of the variable of bathrooms, our model appears to be statistically accurate. All of our diagnostic tests passed, and we have an equation to represent the Airbnb price in the Berkeley/Oakland area, based on number of bedrooms, maximum occupancy, and race.

With regard to the relation between seller's race and asking price, the 95% confidence interval sheds light on our conclusion. The 95% CI here for race is between -0.3673123 and -0.0353176 (Figure 14). Stata's estimated coefficient is -0.2013149, but it is 95% certain that the true value would be captured between -0.3673123 and -0.0353176. When we look at the

race variable (where we used 0 to represent a white seller and 1 to represent an Asian seller), we see that being an Asian seller has a negative effect on rental price. Thus, from the sample of over 100 lodgings being rented on Airbnb, we are 95% confident that being Asian will negatively affect price, assuming that the other factors and variables in our data set are held constant.

lprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
acesq	-.2013149	.0835426	-2.41	0.018	-.3673123	-.0353176

**Figure 14: Final summary of data post-transformation**

## Discussion

After thorough statistical analyses of an Airbnb data set extracted from the Oakland/Berkeley area, we notice that lodging hosts who can be identified as Asian from their profiles tend to have lower rental prices for homes of comparable quality— it turns out hosts believed to be Asian earn much less than their white counterparts even when controlling for other variables related to the property itself.

### Implications

What do our findings suggest about online discrimination and biases toward Asians? If anything, they may support the tacit assumption that Asians must (or may believe they must) exceed expectations in order to receive treatment on the same level as their non-Asian peers. In the realm of online Airbnb transactions, Asians tend to offer a lower price than whites to rent out a lodging of similar quality in a similar neighborhood.

Why are Asian rentals lower? The answer is beyond the scope of this study. Of course, Asians could have found it difficult to get bookings when they offered rental prices comparable to those offered by whites. Alternatively, Asians might offer lower prices to increase the number of booked nights or whites may increase rental prices to reduce the number of bookings. These are ideas for further study.

What can be done? The authors from the Harvard Business School study suggest that Airbnb should consider measures to curb discrimination such as limiting viewership of profile pictures [19]. They argue that there is no need for a potential customer to know a host's name, let alone their appearance, before booking a lodging to stay in. Numerous other e-commerce sites, such as Ebay, prohibit the buyer from seeing any personal information about the seller. Care has to be taken with LinkedIn or Facebook links also.



There will never be a perfect world void of discrimination, and there is bound to be variance among different races. However, our data analysis strongly suggests that there exists, at the very least, an association between Asian American ethnicity and lower rental prices. We support the recommendations of the HBS study that Airbnb should take measures to lessen the likelihood of discrimination.

### Limitations

We also want to acknowledge some of the limitations and shortfalls of our study. First, in our analysis we used the weekly rate, multiplying the daily rate by 7 if no weekly rate was quoted, instead of the daily rate. It is possible that a consumer could receive a discount from the host for renting the lodging for a number of days, making our method a poor way of calculating the true weekly rate. However, as this would have applied to both Asian and white hosts, it would not have affected the overall conclusion of our analysis, only some of the numbers that we used.

Another way we could have added confidence to our analysis would be to account for neighborhood. Although our background research shows that socioeconomic conditions in the Bay Area is diverse, our ZIP codes of interest may have large differences at the neighborhood level. In addition, there are many intangibles that make certain neighborhoods more attractive to renters than others. One possibility for broadening our current study is to account for these variances by assigning a metric to each neighborhood on a scale from 1 to 10. This metric would be based on an algorithm that considers crime rates, survey responses from residents, or ratings from TripAdvisor to rate each neighborhood. Doing this would allow us to have a more multidimensional way of measuring neighborhood quality beyond raw income numbers.

Other limitations in our study may result from other variables that were not included in the statistical modeling having an impact on the result. Examples include gender and age. Another example is the racial mix or preference of the travelers to the Bay Area.

### Next steps

It would be worthwhile to study the normative reasons for price differentials between Asians and whites on Airbnb. Perhaps surveys could be conducted examining how people feel when they encounter an Asian host on Airbnb instead of a white host. Do they feel more at risk in a home owned by an Asian as opposed to one owned by a white host? Do they think homes owned by Asians are generally of lower quality than ones owned by whites? Such surveys would give greater insight into these questions and more if performed.

Further surveys could also investigate the characteristics of hosts, such as the possibility that owners are intentionally underpricing to increase the number of booked nights. Can these findings be replicated in other population segments?

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.  
<http://techscience.org/a/2015090104/>

Beyond paving the way for future studies on racial attitudes toward Asians, this study could also be expanded by testing for biases toward other specific demographics such as homosexuals or females. While it would be relatively easy to test for male/female differences by judging from the profile picture, it would be harder to discern an individual's sexuality from their picture. As the young, loosely regulated sharing economy grows, it will become more and more important to understand who benefits the most and what demographics are limited from participation.

## References

Edelman, B, and Luca M. Digital Discrimination: The Case of Airbnb.com. January 10, 2014. [http://www.hbs.edu/faculty/Publication%20Files/14-054\\_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf](http://www.hbs.edu/faculty/Publication%20Files/14-054_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf)

Cavna, M. 'NOBODY KNOWS YOU'RE A DOG': As iconic Internet cartoon turns 20, creator Peter Steiner knows the idea is as relevant as ever. *The Washington Post*. August 22, 2015. [http://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb\\_blog.html](http://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb_blog.html)

Sweeney, L. Discrimination in Online Ad Delivery. January 28, 2013.  
<http://arxiv.org/ftp/arxiv/papers/1301/1301.6822.pdf>

Edelman, B, and Luca M. Digital Discrimination: The Case of Airbnb.com. January 10, 2014. [http://www.hbs.edu/faculty/Publication%20Files/14-054\\_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf](http://www.hbs.edu/faculty/Publication%20Files/14-054_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf)

Love, D. Blacks' Income Remains the Lowest of Any Group. *TheGrio*. September 18, 2013.  
<http://thegrio.com/2013/09/18/blacks-income-remains-the-lowest-of-any-group/>

Love, D. State of Black America Report: Blacks Are Treated 72 Percent Equal to Whites. *TheGrio*. March 19, 2015. <http://thegrio.com/2015/03/19/urban-league-state-of-black-america-report-2015/>

DeNavas-Walt, C, Proctor B, and Smith J. Income, Poverty, and Health Insurance Coverage in the United States: 2012. U.S Census Bureau. September 2013.  
<http://www.census.gov/2010census/popmap/ipmtext.php?fl=06:0653000>

Tyson, C. ACT's Annual Score Report Shows Languishing Racial Gaps, Mediocre Scores. *InsideHigherEd*. August 20, 2014.  
<https://www.insidehighered.com/news/2014/08/20/acts-annual-score-report-shows-languishing-racial-gaps-mediocre-scores>

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.  
<http://techscience.org/a/2015090104/>

Ramirez, R. Ivy League Schools Don't Reflect U.S. Minority Ratios. *National Journal*. October 19, 2012. <http://www.nationaljournal.com/thenextamerica/education/ivy-league-schools-don-t-reflect-u-s-minority-ratios-20121019>

Pew Research Center. The Rise of Asian Americans. Pew Research Centers Social Demographic Trends Project RSS. April 4, 2013.  
<http://www.pewsocialtrends.org/2012/06/19/the-rise-of-asian-americans/>

Leadership Education for Asian Pacifics, Inc. 2012 API Representation on Fortune 500 Boards. September 2012.  
[https://d3n8a8pro7vhmx.cloudfront.net/leap/pages/75/attachments/original/1413571196/2012\\_LEAP\\_F500\\_FullReport.pdf?1413571196](https://d3n8a8pro7vhmx.cloudfront.net/leap/pages/75/attachments/original/1413571196/2012_LEAP_F500_FullReport.pdf?1413571196)

Berdahl, J, and Min J. Prescriptive Stereotypes and Workplace Consequences for East Asians in North America. American Psychological Association. 2012. <http://www-2.rotman.utoronto.ca/facbios/file/Berdahl%20&%20Min%20CDEMP%202012.pdf>

Mounk, Y. Is Harvard Unfair to Asian-Americans? *The New York Times*. November 24, 2014.  
<http://www.nytimes.com/2014/11/25/opinion/is-harvard-unfair-to-asian-americans.html>

Airbnb. What is Airbnb? <https://www.airbnb.com/help/article/1>

U.S Census Bureau. USA QuickFacts from the US Census Bureau. March 31, 2015  
<http://quickfacts.census.gov/qfd/states/00000.html>

U.S. Census Bureau. 2010 Census Interactive Population Search.  
<http://www.census.gov/2010census/popmap/ipmtext.php?fl=06%3A0653000>

City of Berkeley California. Comprehensive Annual Financial Report for Year Ended June 30, 2014. June 30, 2013.  
[http://www.ci.berkeley.ca.us/uploadedFiles/Finance/Level\\_3\\_-\\_General/FY2014%20Complete%20CAFR.pdf](http://www.ci.berkeley.ca.us/uploadedFiles/Finance/Level_3_-_General/FY2014%20Complete%20CAFR.pdf)

Advameg, Inc. Berkeley, California (CA) Income Map, Earnings Map, and Wages Data.  
<http://www.city-data.com/income/income-Berkeley-California.html>

Edelman, B, and Luca M. Digital Discrimination: The Case of Airbnb.com. January 10, 2014. [http://www.hbs.edu/faculty/Publication%20Files/14-054\\_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf](http://www.hbs.edu/faculty/Publication%20Files/14-054_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf)

---

## Authors

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015. <http://techscience.org/a/2015090104/>

David Wang is a junior at Harvard College concentrating in Statistics with a secondary in Global Health and Health Policy. He is interested in using data analytics to impact healthcare, having worked with various related organizations ranging from U.S. Senator Sherrod Brown with Medicare policy to InCrowd with their pharmaceutical statistics. David enjoys lifting weights and playing sports in his free time.

Stephen Xi is a junior at Harvard College concentrating in Government with a secondary in Computer Science. Academically, he is interested in political institutions and economic development, as well the intersection between technology and politics. Outside of class, Stephen is heavily involved in Harvard Student Agencies and plays violin in the River Charles Ensemble.

John Gilheany is a junior at Harvard College majoring in Statistics with a secondary in Energy & Environment. During his time at Harvard he has done research at HBS and the NASA Tournament Lab. This past summer John worked at Duff & Phelps in Frankfurt, where he stayed in an apartment listed on Airbnb, and at the Brunswick Group in NYC. In his free time, he loves to play basketball and catch up on current events, especially in the technology and energy sectors.

**Editor:** Latanya Sweeney

---

## Citation

Wang D, Xi S, Gilheany J. The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. *Technology Science*. 2015090104. September 1, 2015.

---

## Data

Wang D, Xi S, Gilheany J. Replication Data for: The Model Minority? Not on Airbnb.com. Harvard Dataverse. August 31, 2015. <http://dx.doi.org/10.7910/DVN/UBDYOO>