

Exploring The Importance Of Online Reviews On Price

Qiuyi Chen

Business Analytics, Bentley University
Waltham, MA 02452, U.S.A.

Jiaye Tang

Marketing Analytics, Bentley University
Waltham, MA 02452, U.S.A.

Muhammad Furqan

McCallum Graduate School of Business, Bentley University
Waltham, MA 02452, U.S.A.

Paul D. Berger

Marketing Department, Bentley University
Waltham, MA 02452, U.S.A.

ABSTRACT

This paper considers the importance of online reviews and their effects on different parameters of a business' product or service. We use Airbnb as our source of data and tested customer reviews and other variables within the business to determine their influence on Airbnb's price. We performed our study primarily by using stepwise regression analyses. The results of our analysis are potentially helpful for businesses to develop their marketing and pricing strategies based on how significant variables are.

Keywords: Airbnb, number of reviews, overall satisfaction score, stepwise regression analysis, pricing strategy

INTRODUCTION

In the modern age of marketing, every business gets responses to its products in many ways. Response from buyers in a review form is very important for businesses to evaluate their customer satisfaction, and quality and performance of their product or service. Online reviews are an integral part of a product and service nowadays, and without question it is a pillar in E-Marketing. Customers review the product/service they are considering buying and also provide reviews, telling other people how good or how bad the product/service is. Reviews are treated very seriously in most organizations, and effective responses to the reviews are made in those organizations which know how to treat their customers well. A study found that 30% of shoppers under the age of 45 consult reviews, when possible, for every purchase they make, while 86% of these shoppers say that reviews are essential in making purchase decisions (Maslouska et al. 2016).

Effective marketing strategies can usually be developed if we know what drives consumer behavior to choose or not to choose a certain product or service. For example, if a consumer looks up reviews of a particular brand of sneakers, would he/she buy that brand even if they haven't tried it themselves? A web-based experiment found that after reading reviews, those exposed to the highly positive (vs. negative) reviews evaluated the product more positively, which, in turn, led to a stronger purchase intention (Lee et Shin, 2014).

A key question arises: if we have more positive reviews of a product, what is our next marketing or pricing strategy? For this purpose, we are studying the relationship between price and number and type of online reviews. If a product has more positive online reviews, would its price be higher than the same kind of product with fewer positive reviews? A lot of different researchers have worked previously on the relationship between consumer behavior and reviews. We consider getting a sense of what consumers will be willing to pay if their product has more positive or fewer positive reviews. To study this issue, we selected Airbnb, with more than half a billion guest arrivals to date, and accessible in 62 languages across 191 countries and regions. Airbnb indicates that it promotes people-to-people connections, community, and trust around the world. (Airbnb, 2019).

We discuss in this paper how the number of, and satisfaction expressed in, reviews affect price. Most companies offer scores on a 1-5 set of scale options. That score can help predict what a company's pricing strategy should be after evaluation of the relationship between price and review score.

LITERATURE REVIEW

In 2016, Askalidis and Malthouse found that as the number of reviews of a product increase, the conversion rate escalates. For having five reviews, the likelihood of purchasing increases nearly 300% on average, compared to having no reviews (Askalidis and Malthouse, 2016). They have also found that the value of reviews on electronics products is the highest (an increase of 317%) followed by that of household products and Apparel. Furthermore, the value of reviews plays a more important role on higher-priced items. This result is predictable; as the price gets higher, there are more risks involved in the purchase process. Therefore, customers want to do more research concerning the potential purchase. In such a situation, having reviews help customers to mitigate risks.

In 2016, Kostyra et al. found that three dimensions of online customer reviews (valence, volume, and variance) affect customer choice: keeping all other variables constant, the higher the valence, the higher the likelihood of customer choice. Volume and variance moderate the impact of the valence on customer choice. For example, the likelihood of customer choice rises with the number of positive reviews. However, the number of reviews has minimal effects if the reviews are neutral or negative. For the third dimension of online customer reviews, the higher the variance, customers' likelihood of purchase increases for products/services with a low-level valence, but decreases for products/services with high-level or medium-level valence. Uncertainty, which is represented by high variance, benefits poorly-rated products/services and harms products/services that have more highly-positive reviews (Kostyra et al., 2016).

METHODOLOGY

The data used for our study come from *Airbnb Collection Data*, which is an independent, non-commercial set of tools to allow people to explore the uses of Airbnb in the real-world. This website directly sourced the information from the Airbnb site and it includes data from five continents and around a hundred and five cities. For this paper's purpose, which is to investigate the effect of reviews on price, we chose Boston Airbnb's dataset from September 2016 as our main resource. This dataset serves to provide a spectrum of aspects that customers may be curious about before booking an Airbnb room. Specifically, this dataset includes the number of reviews, overall satisfaction rate of the reviews, "accommodations" (the dataset's phrase for the number of beds), type of housing situation, the number of bedrooms, bathrooms, the minimum stay of nights required, longitude, latitude, and the different locations of Boston from Chinatown (extreme East) to Allston (extreme West), twenty-five locations in total.

The only dependent variable in this study is *price* [per night] and the primary independent variables that we wish to study in this paper are the number of reviews and the “quality” (positivity vs. negativity) of the reviews. The latter is scaled from 1-5 and referred to as the “overall satisfaction score.” As noted, other variables, such as the house type and the number of rooms, are also studied. The two sets of dummy variables in our dataset are the house type and the location. For house type, we have three groups which are the entire house, private room, or shared room. For the location, it includes 25 groups: Allston, BackBay, Bay Village, Beacon Hill, Brighton, Charlestown, Chinatown, Dorchester, Downtown, East Boston, Fenway, Hyde Park, Jamaica Plain, Leather District, Longwood Medical Area, Mattapan, Mission Hill, North End, Roslindale, Roxbury, South Boston, South Boston Waterfront, South End, West End, and West Roxbury.

We will first explore the stepwise regression results between price and all variables, excluding reviews and the locations (i.e., housing arrangement and room features only.) Next, we conduct a stepwise regression analysis between price and all variables, excluding only location; this includes the two aspects of reviews, number and satisfaction score. For the final phase, we include all the variables, including location; we do this in two steps, one using longitude and latitude only, and the other using the specific 25 locations.

Hypotheses

Based on a conceptual framework of our literature review as well as personal experience booking Airbnb rooms, we hypothesized the following for the stepwise regression results:

1. Assuming other variables in the model are held constant, the higher the number of reviews, the higher the price.
2. Assuming other variables in the model are held constant, the higher the overall satisfaction score, the higher the price.
3. Assuming other variables in the model are held constant, the higher the number of rooms (bedrooms or bathrooms), the higher the price.
4. Assuming other variables in the model are held constant, the better the housing arrangement (entire home > private room > shared room), the higher the price.
5. Assuming other variables in the model are held constant, the higher the number of accommodations (i.e., number of beds), the higher the price.
6. Assuming other variables in the model are held constant, the larger the number of the required minimum stay-nights, the *lower* the price.

ANALYSIS

As mentioned above, we will run four stepwise regressions. The first stepwise regression contains housing/room features only. Then, we will run a stepwise regression with housing/room features and customer reviews – the two aspects, number of and satisfaction level of. Next, we will run stepwise regressions with housing/room features, customer reviews, and location. This will consist of two analyses, one representing location by longitude and latitude, the other representing location by name of the area (e.g., “downtown.”) The four stepwise regressions are summarized as follows:

- 1) $Y = \text{price}$, $X = \text{basic housing/room characteristics}$
- 2) $Y = \text{price}$, $X = \text{basic housing/room characteristics} + \text{customer reviews}$
- 3a) $Y = \text{price}$, $X = \text{basic housing/room characteristics} + \text{customer reviews} + \text{location represented by longitude and latitude}$
- 3b) $Y = \text{price}$, $X = \text{basic housing/room characteristics} + \text{customer reviews} + \text{location represented by location names (e.g., “East Boston”)}$

STEPWISE REGRESSION WITH HOUSING/ROOM CHARACTERISTICS ONLY

The results show that the number of rooms and the housing/room type (Entire home/apt) are significant. If it is an entire home/apartment, the price will increase by about 94 dollars (see 2nd table below.) An entire home/apartment has more space and can provide indoor activities, which explains why the price is higher.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.596a	0.356	0.355	77.861
2	.686b	0.47	0.469	70.642
3	.700c	0.49	0.489	69.317
4	.704d	0.496	0.495	68.912

a Predictors: (Constant), Entirehomeapt

b Predictors: (Constant), Entirehomeapt, bedrooms

c Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms

d Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms, accommodations

Model		Unstandardize d Coefficients	Std. Error	Standardized Coefficients Beta	t	Sig.
4	(Constant)	7.185	5.222		1.376	0.169
	Entirehomeapt	94.469	3.875	0.48	24.381	0
	bedrooms	30.007	3.412	0.216	8.796	0
	bathrooms	34.552	4.254	0.146	8.122	0
	accommodations	7.053	1.502	0.126	4.697	0

a Dependent Variable: price

For the number of bedrooms, bathrooms, and accommodations (beds), it is not surprising that they are significant. With all other variables held constant, one more bedroom will increase the price by about 30 dollars. One more bathroom will increase the price by about 34 dollars. It seems like the price for a bathroom is higher than that of bedrooms. With all other variables held constant, one more accommodation (bed) will increase the price about 7 dollars.

STEPWISE REGRESSION WITH REVIEWS AND OTHER VARIABLES, EXCEPT LOCATION

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.596a	0.356	0.355	77.861
2	.686b	0.47	0.469	70.642
3	.700c	0.49	0.489	69.317
4	.704d	0.496	0.495	68.912
5	.708e	0.501	0.499	68.622
6	.710f	0.505	0.503	68.358

a Predictors: (Constant), Entirehomeapt

b Predictors: (Constant), Entirehomeapt, bedrooms

c Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms

d Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms, accommodations

e Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms, accommodations, reviews

f Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms, accommodations, reviews, overall_satisfaction

Model		Unstandardized Coefficients	Std. Error	Standardized Coefficients	t	Sig.
6	(Constant)	-60.339	19.771		-3.052	0.002
	Entirehomeapt	91.891	3.872	0.467	23.731	0
	bedrooms	29.243	3.387	0.211	8.634	0
	bathrooms	34.495	4.22	0.146	8.175	0
	accommodations	7.589	1.493	0.135	5.083	0
	reviews	-0.175	0.041	-0.072	-4.284	0
	overall_satisfaction	15.818	4.106	0.065	3.853	0

a Dependent Variable: price

After adding number of reviews and satisfaction scores, the R square increased to 0.506 from 0.496, and both variables were significant. According to the above table, the result of the number of reviews is different from our hypothesis. When the number of reviews increases, the price will slightly decrease.

Model Summary

Model	R	R Square	Adjusted R S	Std. Error of the Estimate
1	.064a	0.004	0.004	0.3949

a Predictors: (Constant), reviews

Model	Unstandardized Coefficient B	Std. Error	Standardized Beta	t	Sig.
1 (Constant)	4.637	0.012		397.619	0
reviews	0.001	0	0.064	2.704	0.007

a Dependent Variable: overall_satisfaction

There are potentially two explanations for this. First, when then an Airbnb house receives more reviews, it is likely that it will get more *negative* review scores (as well as positive ones), and this could cause the price to be lower. We applied a linear regression analysis to test this assumption (see tables immediately below.) However, when the number of reviews increased, the overall satisfaction *increased*. From the results below, the number of reviews only explains .004 of the overall satisfaction, although it is statistically significant. And when the number of reviews increases, the price will slightly increase.

However, a lower price will attract more customers, and that often leads to an increased number of reviews. To examine this, we tested price vs. number of reviews. The results below show that, indeed, when price decreases, the number of reviews will increase (exemplified by the negative coefficient of *price* in the second table below.) Of course, this likely just reflects the increased number of patrons.

Model Summary

Model	R	R Square	Adjusted R S	Std. Error of the Estimate
1	.128a	0.016	0.016	39.657

a Predictors: (Constant), price

Model	Unstandardized Coefficient B	Std. Error	Standardized Beta	t	Sig.
1 (Constant)	38.26	1.801		21.245	0
price	-0.053	0.01	-0.128	-5.435	0

a Dependent Variable: reviews

Based on this, we have two conclusions. First, when the number of reviews is lower and the average review (satisfaction) score is higher, we can lower our price to get more reservations. Second, all else remaining constant, when we get a higher (average) satisfaction score, we can consider increasing price, to generate increased profit.

STEPWISE REGRESSION WITH REVIEWS AND OTHER VARIABLES, AND LATITUDE AND LONGITUDE

In this stepwise regression, we add latitude and longitude; the north and east part of Boston being indicated as more expensive.

Model	Unstandardized Coefficients		Standardized Coefficient	Sig.
	B	Std. Error		
8 (Constant)	-2815.062	5587.97		0.614
Entirehomeapt	80.276	3.919	0.408	0
bedrooms	30.948	3.299	0.223	0
latitude	511.045	71.195	0.13	0
bathrooms	32.215	4.099	0.137	0
accommodations	8.132	1.449	0.145	0
longitude	265.684	53.2	0.088	0
reviews	-0.201	0.04	-0.083	0
overall_satisfaction	17.831	3.989	0.073	0

a Dependent Variable: price

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.596a	0.356	0.355	77.861
2	.686b	0.47	0.469	70.642
3	.703c	0.494	0.493	69.022
4	.715d	0.512	0.511	67.826
5	.720e	0.519	0.517	67.38
6	.724f	0.524	0.522	67.041
7	.728g	0.53	0.528	66.645
8	.731h	0.535	0.533	66.293

a Predictors: (Constant), Entirehomeapt

b Predictors: (Constant), Entirehomeapt, bedrooms

c Predictors: (Constant), Entirehomeapt, bedrooms, latitude

d Predictors: (Constant), Entirehomeapt, bedrooms, latitude, bathrooms

e Predictors: (Constant), Entirehomeapt, bedrooms, latitude, bathrooms, accommodates

f Predictors: (Constant), Entirehomeapt, bedrooms, latitude, bathrooms, accommodates, longitude

g Predictors: (Constant), Entirehomeapt, bedrooms, latitude, bathrooms, accommodates, longitude, reviews

h Predictors: (Constant), Entirehomeapt, bedrooms, latitude, bathrooms, accommodates, longitude, reviews, overall_satisfaction

STEPWISE REGRESSION WITH REVIEWS AND OTHER VARIABLES, AND SPECIFIC PARTS OF THE CITY OF BOSTON

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
23	.795w	0.632	0.627	59.219

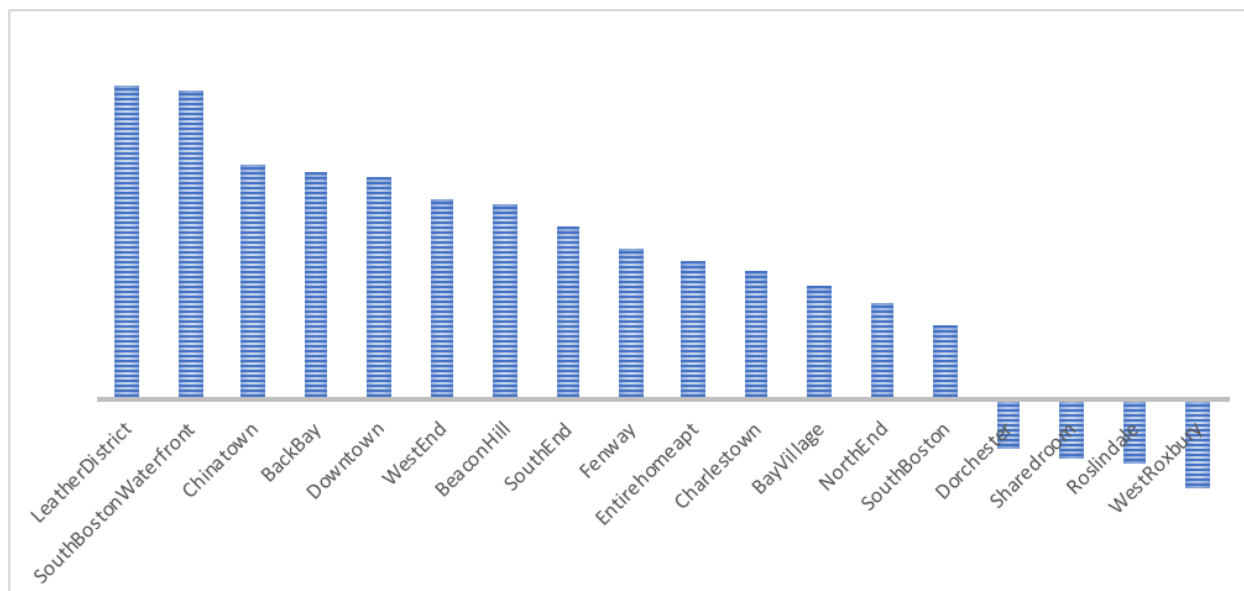
w Predictors: (Constant), Entirehomeapt, bedrooms, bathrooms, BackBay, SouthBostonWaterfront, BeaconHill, Downtown, SouthEnd, Fenway, Chinatown, accommodations, Dorchester, Charlestown, NorthEnd, SouthBoston, WestEnd, LeatherDistrict, overall_satisfaction, reviews, BayVillage, WestRoxbury, Roslindale, Sharedroom

Model		Unstandardized Coefficient			Standardized t	Sig.
		B	Std. Error	Beta		
23	(Constant)	-59.732	17.529		-3.408	0.001
	Entirehomeapt	53.595	3.783	0.272	14.169	0
	bedrooms	42.167	3.015	0.304	13.985	0
	bathrooms	27.519	3.727	0.117	7.383	0
	BackBay	87.387	5.752	0.247	15.191	0
	SouthBostonWaterfr	118.361	10.396	0.171	11.386	0
	BeaconHill	74.925	5.949	0.198	12.595	0
	Downtown	85.363	7.293	0.18	11.705	0
	SouthEnd	66.188	5.474	0.193	12.091	0
	Fenway	57.723	6.388	0.139	9.036	0
	Chinatown	89.905	14.207	0.093	6.328	0
	accommodations	7.671	1.304	0.137	5.882	0
	Dorchester	-18.194	5.544	-0.051	-3.282	0.001
	Charlestown	49.979	9.309	0.08	5.369	0
	NorthEnd	37.433	6.746	0.086	5.549	0
	SouthBoston	29.203	6.389	0.069	4.571	0
	WestEnd	77.11	18.055	0.062	4.271	0
	LeatherDistrict	120.265	34.339	0.051	3.502	0
	overall_satisfaction	11.617	3.611	0.047	3.217	0.001
	reviews	-0.106	0.036	-0.044	-2.962	0.003
BayVillage	44.073	18.894	0.034	2.333	0.02	
WestRoxbury	-33.825	14.557	-0.034	-2.324	0.02	
Roslindale	-24.078	10.302	-0.034	-2.337	0.02	
Sharedroom	-22.362	9.729	-0.034	-2.298	0.022	

a Dependent Variable: price

The most expensive places are the LeatherDistrict, SouthBostonWaterfront, Chinatown, BackBay, Downtown, WestEnd, BeaconHill, Southend Fenway, Charlestown, BayVillage, NorthEnd, SouthBoston, Dorchester, Roslindale, WestRoxbury.

The graph below indicates the positive/negative impact of the top 18 locations in terms of magnitude of effect.



The basic results are not very different for the review variables and other variables in either of the analyses involving location.

CONCLUSIONS

Most of our stepwise regression results agree with our hypotheses while a few of them do not. The following findings support our hypotheses:

1. Holding all other variables constant, the better the housing type, the higher the price. Housing types were defined in the Hypotheses section.
2. Holding all other variables constant, the number of rooms, including bedrooms and bathrooms, as well as the accommodations (number of beds), the higher the price.
3. Holding all the other variables constant, the higher the overall satisfaction scores, the higher the price. This, of course, makes perfect sense.
4. Holding all the other variables constant, the price of the Airbnb room that could be charged is higher for the north and east parts of Boston. In addition, the leather district has the most expensive price for Airbnb Boston locations. The leather district, features 19th-century brick warehouses and has many luxury lofts. To most, it is no surprise that this location charges the highest price among Boston's Airbnb offerings.

The two results that do not support our hypotheses are:

1. Holding all the other variables constant, as the number of reviews increases, the lower the price.
2. The number of minimum stay nights does not have a significant relationship with price (at least incremental to other variables in the stepwise regressions) as it does not show in any stepwise regression result.

As we have discussed earlier, there are two possible explanations for the result of the increasing number of reviews leads to a decreasing price. The first possible explanation is that it is always easier for people to leave comments when they have unsatisfied feelings rather than appreciation. Therefore, when seeing a high number of reviews, this might indicate a lower overall satisfaction score. However, after testing the relationship between the number of reviews and the overall satisfaction score, we found that the R square is only 0.004 which is not significant. Thus, we reject that explanation. The second potential explanation is that people are attracted by lower-priced Airbnb rooms, and since more people choose those rooms, the chance for a review to be left is higher.

For the variable of minimum stay nights, the result shows that price and the number of minimum stay nights do not have a correlation.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

A key limitation of our research is that we evaluated only reviews and overall satisfaction. Our research does not include customers' attitudes, or their reasons for positive or negative reviews toward their living experience. In other words, we did not go into further details of their reviews.

Another key limitation of this research is that we examined only one city (Boston); we do not know if our results would be duplicated for other cities.

Some areas for future research involve the two points made immediately above.

References

- Airbnb. (2019). About us. Retrieved from <https://news.airbnb.com/en-in/about-us/>.
- Askalidis, Georgios & Malthouse, Edward. (2016). The Value of Online Customer Reviews. *ACM Conference on Recommender Systems (RecSys16)*, Boston, MA 10.1145/2959100.2959181.
- Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2016). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing*, 33(1), 11–26. doi: 10.1016/j.ijresmar.2014.12.004
- Lee, Eun-Ju & Shin, Soo Yun. (2014). When do consumers buy online product reviews? Effects of review quality, product type, and reviewer's photo. *Computers in Human Behavior*. 31. 356–366. 10.1016/j.chb.2013.10.050.
- Maslowska, E., Malthouse, E. and T. Collinger (2016). The Customer Engagement Ecosystem. *Journal of Marketing Management*, 32(5–6), 469–501.