ISSN: 2394-7659
IMPACT FACTOR- 2.789

International Journal of Engineering Researches and Management Studies

ESTIMATING ROOM RENTS FOR VARIOUS HOTELS ACROSS INDIA

Lalit Kumar Aggarwal[*1], Vaibhav Raj[2] & Nav Krishna Choudhary[3]
[*1,2&3]Indian Institute of Management Lucknow. Prabandh Nagar– 226013, Uttar Pradesh, INDIA

ABSTRACT
The paper exposes the economic viability of the Indian Hotel Industry by employing primary data taken from Hotels.com website. The paper focusses to build a relationship between hotel prices and various explanatory variables like hotel star rating, city population and distance from the airport. Apart from these regressors we used some dummy variables like availability of swimming pool, whether free wifi available, free breakfast, whether city is a metro city, whether city is a tourist destination and is the day weekend. After our analysis we found that whether the city is metro city & is the day a weekend comes out to be insignificant variables. As we are using more than 10 regressors, we have a problem of multicollinearity. Also, the data is taken from hotels from 18th Dec 2016 to 08th Jan 2017. There can be a problem of Autocorrelation but we found an insignificant relation as the time series data is only for 8 days. On the whole, our study shows that hotel rate in India is dependent mainly on traditional factors like star ratings and availability of breakfast and wifi..

KEYWORDS: Tourism, Hotels, Hotel rents in India.

1. INTRODUCTION

India, after China, is considered as one of the most lucrative hotel markets in the world. There are multiple factors that are driving the growth of the hospitality sector. The potential for economic growth, increase in disposable incomes and the growing middle class are expected to drive both leisure and business travel in the coming years.

The hotel industry in India thrives because of the growth in tourism and travel. Due to the increase in tourism with rising foreign and domestic tourists, hotel sector is bound to grow. There is an emergence of budget hotels in India to cater to the majority of the population who seek affordable stay. International companies are also increasingly looking at setting up such hotels. Imbalance in increase in tourists both domestic and foreign not been supported with equal number of rooms is a latent source of opportunity for growth.

Tourism is the third largest foreign exchange generator of the country. The booming tourism industry has had a cascading effect on the hospitality sector. The hospitality sector has resulted in an increase in the occupancy ratios and average room rates. The long term outlook for the Indian hospitality business continues to be positive, both for the business and leisure segments with the potential for economic growth, increases in disposable incomes and the burgeoning middle class.

This report analyses that how independent factors like star rating, distance from airport, population of city, free Wi-Fi, free breakfast, etc affect the hotel rent in several cities of India.

The analysis contains the relationship between the hotel rent to the independent factors individually. It also includes a regression model with all the independent factors taken together, as well as a regression model with few of the independent factors taken together on the pricing of hotels. For example, the impact of star rating, free breakfast and free Wi-Fi taken cumulatively on the hotel pricing.

Parameters like Star Rating, New Year Eve, Metro City, Tourist Destination, distance from airport, free Wi-Fi, free breakfast and swimming pool have a positive correlation with room rent. Whereas, city population and hotel capacity have a negative correlation with room rent. The parameter ‘Is Weekend’ is negatively correlated with room rent

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

2. DATA ANALYSIS

The data we took for our analysis was a panel data. The rents corresponding to all the hotels are taken for 8 different dates from 18th Dec 2016 to 8th Jan 2017. The motivation to choose such date was to analyze the effect on the hotel prices on festive seasons like Christmas and New Years eve. All the data i.e. hotel rent is collected from the same website as to remove any discrepancy from other sites. The site used to extract the data is www.hotels.com.

The extraction of the data manually was a time taking exercise. Hence, we took help from Manickavelan (PGP32183). Manickavelan wrote a code which directly extracted all the data from the concerned 37 cities across India for all the dates. The snapshot of the user interface for the code is shown in Fig. 1 below.

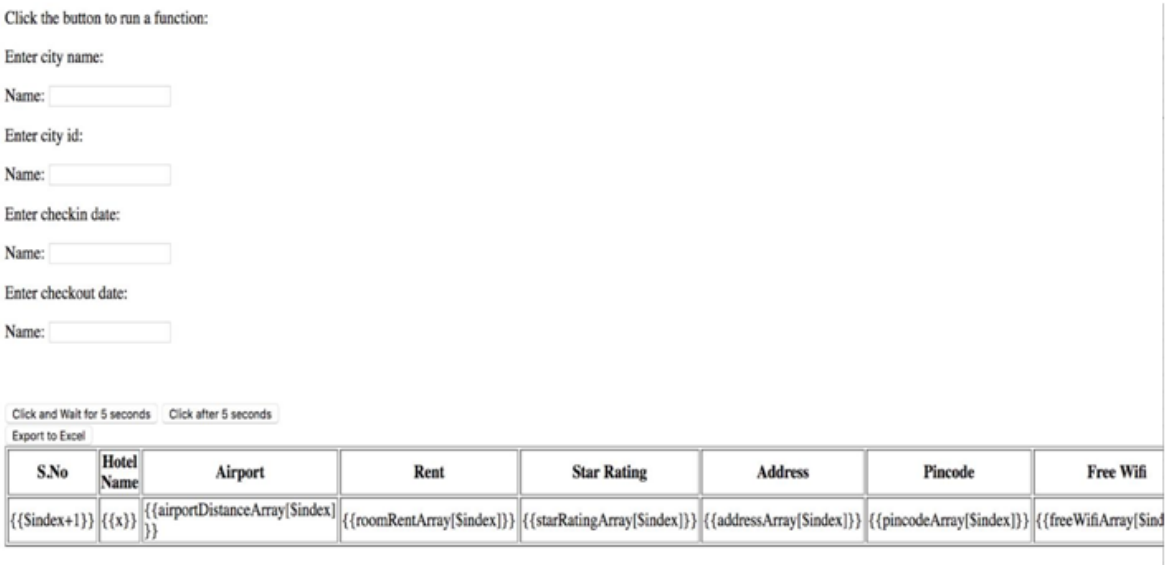


Fig. 1: User-Interface of the code to extract the data from website – hotels.com

3. METHODOLOGY

Based on data from 2016 hotels across 9 cities from various parts of India, we have analysed the impact of various parameters on the hotel prices. The primary objective is to find a model which has the best fit between the independent variables and the dependent variable (in this case, price). The objectives of our study are:

- Analysis on how individual factors have a bearing on the pricing of hotels. For example, the impact of star rating of a hotel on its pricing, the impact of the distance from the airport on the pricing
- Analysis on how all the independent factors taken combined have an impact on the pricing of hotel rooms
- Analysis on the impact of a combination of a few independent factors taken together on the pricing of hotels. For example, the impact of star rating, free breakfast and free wi-fi taken cumulatively on the hotel pricing

Room Rent = β_0 + β_1*Star Rating [1] + β_2*Is New Year Eve + β_3*Is Weekend + β_4*Is Metro City + β_5*Is Tourist Destination + β_6*City Population + β_7*Airport + β_8*Free Wifi + β_9*Free Breakfast + β_{10}*Hotel Capacity + β_{11}*Has Swimming Pool + β_{12}*Star Rating [2] + β_{13}*Star Rating [3] + β_{14}*Star Rating [3] + β_{15}*Star Rating [4]

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

This above equation is the initial equation that we assumed. It is a linear equation in both variables and parameter and has one constant and 11 explanatory variables. Star Rating ranges only from 1 to 5, that's why we use 4 different dummy variables to explain 5-star rating categories.

Following are the explained and explanatory variables:

- Hotel Rent – One-day price for every hotel
- Star Rating – The rating given by hotels.com to each hotels. This is constant across time. Star rating takes only 5 values i.e. 1, 2, 3,4 & 5. Thus, we have to take 4 different dummy variables.
- New Years Eve – Dummy variable with a value 1 if day is 31st Dec 2016
- Is Weekend – Dummy variable with a value 1 in case the day is Saturday & Sunday
- Is Tourist Destination – Dummy variable with a value 1 if the city is a tourist destination
- City Population – Population of the city
- Airport – Hotel's distance from airport
- Free Wifi – Dummy variable with a value 1 if free wifi is available in hotel
- Free Breakfast – Dummy variable with a value 1 if free wifi is available in hotel
- Has Swimming Pool – Dummy variable with a value 1 if swimming pool is available in hotel

4. THEORITICAL BACKGROUND

Tourism is an important part of economy of India. There are many tourist places of varying importance– some for historical reasons, some for religious regions, and others for their modernity. India earned about INR 135193 crore INR equivalent of foreign exchange from tourism in 2015. While there are many foreign visitors to India, due to its size internal tourism is more important (8.03 million foreign arrivals versus 1432 domestic visits in 2015). We explore a supply side effect, as to how hotel rooms are priced. For this, we collected prices of hotels in 9 cities – Agra, Ahmedabad, Delhi, Goa, Hyderabad, Jaipur, Kolkata, Lucknow, Pune – on 8 days (between 18th December and 8th January). In total we have collected data for 793 hotel

Table 1: Result of the first regression without checking for Multi-collinearity, Heteroskedasticity and Serial correlation

Variable	Coefficient	Std. Error	t-Statistic	p value
AIRPORT	45.841	1.14E+01	4.024463	0.00
FREEBREAKFAST	-458.3578	190.84	-2.401791	0.02
FREEWIFI	952.3771	400.5655	2.377581	0.02
HASSWIMMINGPOOL	2204.771	258.2648	8.536864	-
HOTELCAPACITY	-13.29711	1.561907	-8.51338	-
ISMETROCITY	-566.8143	306.715	-1.848016	0.06
ISNEWYEAREVE	646.31	281.7646	2.293794	0.02
ISTOURISTDESTINATION	1398.046	201.8089	6.927571	-
ISWEEKEND	-126.871	192.3672	-0.659525	0.51
Star Rating [2]	2146.582	2541.232	0.844701	0.40
Star Rating [3]	4935.782	2523.034	1.956288	0.05
Star Rating [4]	3368.51	2515.165	1.33928	0.18
Star Rating [5]	11520.86	2543.456	4.529611	-
POPULATION	-0.000139	4.59E-05	-3.032709	0.00
CONSTANT	-1223.157	2556.877	-0.478379	0.63

5. ECONOMETRIC TESTS

The biggest problem that we thought while running the regression is that there are too many regressor (explanatory variables) which is the root for the multi-collinearity problem. To check for multi-collinearity, we have used the below mentioned 2 test:

- Correlation Table
- VIF (Variance Inflation Factor)

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

For the problem of Heteroskedasticity, following tests are performed to check the constant variance of the data:

- Residual Plot
- White Test
- Breusch Pagen Godfrey Test

Lastly, as we have taken a panel data, we have to check for the serial correlation (autocorrelation). Following are the test performed to check the serial correlation of the data:

- Breusch Godfrey Serial Correlation LM Test

Moreover, the time series data is only for 8 days, thus we don't have any problem of any structural break

6. RESULTS

Test for Multi-collinearity

Correlation Matrix

Table 2: Correlation Matrix

Correlation	ROOMRENT	POPULATION	ISWEEKEND	STDESTINATION	ISNEWYEAREVE	ISMETROCITY	MMINGPOOL	FREEWIFI	FREEBREAKFAST	AIRPORT	STARRATING	HOTELCAPACITY
ROOMRENT	1.00	-0.04	-0.00	0.15	0.03	0.02	0.31	0.00	-0.03	0.03	0.34	0.17
POPULATION	-0.04	1.00	0.00	-0.06	-0.00	0.63	0.05	0.15	0.18	0.22	0.19	0.28
ISWEEKEND	-0.00	0.00	1.00	-0.00	0.29	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00
ISTOURISTDESTINATION	0.15	-0.06	-0.00	1.00	0.00	0.29	0.23	-0.12	-0.11	-0.02	0.12	0.06
ISNEWYEAREVE	0.03	-0.00	0.29	0.00	1.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00
ISMETROCITY	0.02	0.63	-0.00	0.29	-0.00	1.00	0.18	0.03	0.04	0.04	0.14	0.17
HASSWIMMINGPOOL	0.31	0.05	0.00	0.23	0.00	0.18	1.00	-0.02	0.01	0.00	0.65	0.50
FREEWIFI	0.00	0.15	-0.00	-0.12	0.00	0.03	-0.02	1.00	0.17	0.00	0.01	-0.01
FREEBREAKFAST	-0.03	0.18	-0.00	-0.11	0.00	0.04	0.01	0.17	1.00	0.01	0.05	0.02
AIRPORT	0.03	0.22	-0.00	-0.02	0.00	0.04	0.00	0.00	0.01	1.00	-0.06	-0.00
STARRATING	0.34	0.19	0.00	0.12	-0.00	0.14	0.65	0.01	0.05	-0.06	1.00	0.64
HOTELCAPACITY	0.17	0.28	0.00	0.06	0.00	0.17	0.50	-0.01	0.02	-0.00	0.64	1.00

Correlation matrix shows that most of the variables not correlated.

Only the following 3 significant combination showed a correlation of ~0.65

- Population –Is Metro City
- Star Rating – Swimming Pool
- Hotel Capacity – Star Rating

We choose to do nothing as removing regressor will create Specification bias

VIF (Variance Inflation Factor)

Table 3: VIF Table

	Auxillary R square	VIF	Comment	Multicollinearity
Airport	0.1121	1.13	Less than 5	No
Hasswimmingpool	0.5081	2.03	Less than 5	No
Freewifi	0.0609	1.06	Less than 5	No
FreeBreakfast	0.0695	1.07	Less than 5	No
Hotel Capacity	0.5221	2.09	Less than 5	No
IsMetroCity	0.5409	2.18	Less than 5	No
IsNewYearEve	0.0851	1.09	Less than 5	No
Population	0.5785	2.37	Less than 5	No

ISSN: 2394-7659
IMPACT FACTOR- 2.789

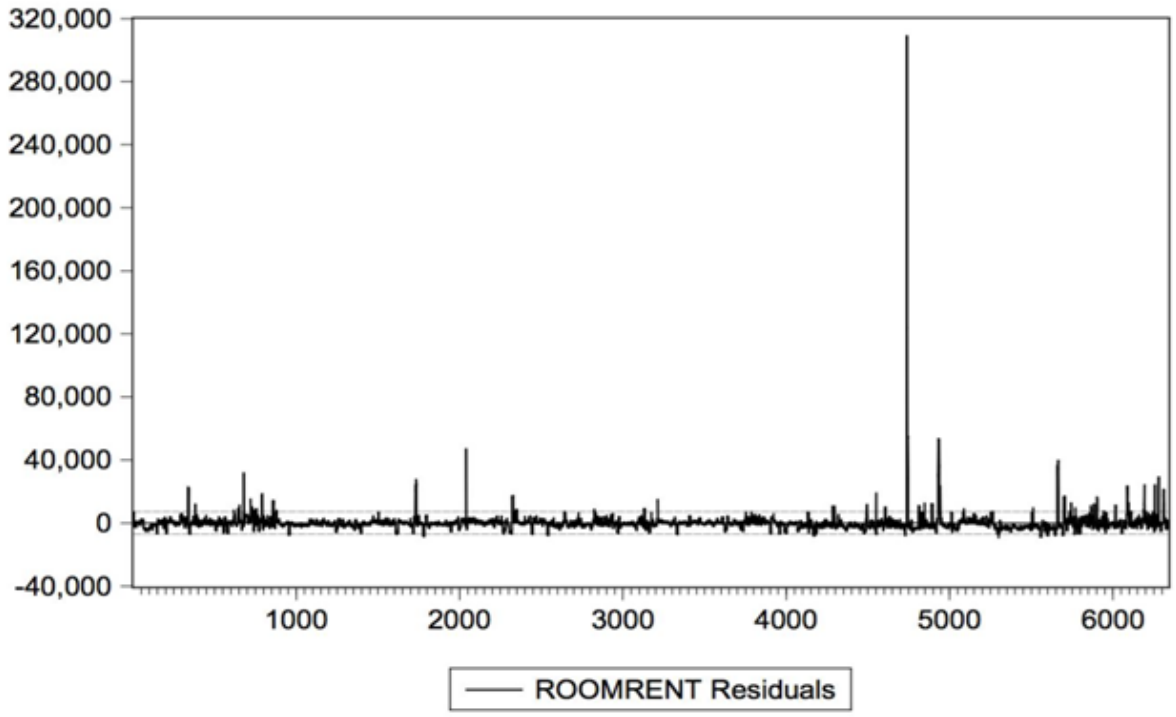
IJERMS
International Journal of Engineering Researches and Management Studies

The VIF for all the regressor is less than 5 and thus we can conclude that there is no multi collinearity in the data used for regression.

Test for Heteroskedasticity

Residual Plot

From the above 2 test, we conclude that there is not severe multicollinearity in the data.



Comment: The residual plot shows that there is one outlier in the total data sets. The outlier is a hotel from Rajasthan that charges a rent of Rs 300,000 per night. If we remove the outlier and observe the residual plot we can see that there is no such case of Heteroskedasticity.

White Test

In statistics, the White test is a statistical test that establishes whether the variance of the errors in a regression model is constant: that is for homoskedasticity.This test, and an estimator for heteroskedasticity-consistent standard errors, were proposed by Halbert White in 1980. These methods have become extremely widely used, making this paper one of the most cited articles in economics

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

Table 4: White Test Result

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 08/21/17 Time: 18:38
Sample: 1 6338
Included observations: 6337
Collinear test regressors dropped from specification

F-statistic	2.209534	Prob. F(93,6243)	0
Obs*R-squared	201.9341	Prob. Chi-Square(93)	0
Scaled explained SS	117014	Prob. Chi-Square(93)	0

R-squared	0.031866	Mean dependent var	50109311
Adjusted R-squared	0.017444	S.D. dependent var	1.71E+09
S.E. of regression	1.70E+09	Akaike info criterion	45.35459
Sum squared resid	1.79E+22	Schwarz criterion	45.45478
Log likelihood	-143612	Hannan-Quinn criter.	45.38929
F-statistic	2.209534	Durbin-Watson stat	0.856886
Prob(F-statistic)	0		

Breusch Pagen Godfrey Test

Table 5: Breusch Pagan Godfrey Test

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 08/21/17 Time: 18:39
Sample: 1 6338
Included observations: 6337

F-statistic	3.554349	Prob. F(14,6322)	0
Obs*R-squared	49.48942	Prob. Chi-Square(14)	0
Scaled explained SS	28677.45	Prob. Chi-Square(14)	0

R-squared	0.00781	Mean dependent var	50109311
Adjusted R-squared	0.005612	S.D. dependent var	1.71E+09
S.E. of regression	1.71E+09	Akaike info criterion	45.3542
Sum squared resid	1.84E+22	Schwarz criterion	45.37019
Log likelihood	-143690	Hannan-Quinn criter.	45.35974
F-statistic	3.554349	Durbin-Watson stat	0.84914
Prob(F-statistic)	0.000007		

Comment: Both the White Test and Breusch Pagan Godfrey Test have p-value tending to zero. We reject the null hypothesis for both the test at 99% level of significance. Thus, we can conclude that there is Heteroskedasticity in our data.

Correction: We use two correction measures for removing the heteroskedasticity from our data. Following are the 2 correction measures:

- ***Take logarithm to population*** – This is because population cannot be directly proportional to the house rent. Moreover, there would be a city with a minimum population yet charging some rent. Hence, taking log will curb such problem.

Robust Standard Error – To solve the problem of Heteroskedasticity, we take robust standard error

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

Test for Autocorrelation

Breusch Godfrey Serial Correlation LM Test

Table 6: Breusch Godfrey Serial Correlation Lm Test Results

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	3335.572	Prob. F(2,6320)	0.078
Obs*R-squared	3254.143	Prob. Chi-Square(2)	0.069

Variable	Std. Error	p-value
POPULATION	3.20E-05	0.94
Star Rating [2]	1772.788	0.87
Star Rating [3]	1754.599	0.89
Star Rating [4]	1760.09	0.89
Star Rating [5]	1774.361	0.82
ISWEEKEND	134.1957	0.87
ISTOURISTDESTINATION	140.7811	0.91
ISNEWYEAREVE	196.558	0.94
ISMETROCITY	213.9634	0.98
HOTELCAPACITY	1.089615	0.82
HASSWIMMINGPOOL	180.1668	0.83
FREEWIFI	279.4328	0.98
FREEBREAKFAST	133.134	0.94
AIRPORT	7.946053	0.82
C	1783.703	0.86

The data that we took has no autocorrelation according to Breusch Godfrey Serial Correlation LM Test.

p-value is 0.078 which is not significant at level of significance of 95%.

Comment – There was no autocorrelation in our data. P -value is greater than 5%. Also, the data is for only 8 different days. Thus, theoretically also, there should not be any autocorrelation and the result concretes it.

Corrections – Following are the corrections we took to improve our regression:

- Took log for population data
- Removed insignificant variable – Is Metro City & Is Weekend
- Took Robust Standard Error to remove Heteroskedasticity

Final Result

The final regression shows that a hotel at an average charges Rs 1195 if the hotel is located in a tourist destination. Also, having swimming pool and free wifi has a positive correlation with the hotel rent. But, contrastly, free breakfast has a negative correlation with the hotel rent. This relation is difficult the explain theoretically and thus we can say that is is a spurious correlation and not a causal relation. One more thing that is astonishing is that we have a positive correlation of the hotel rent with the distance from airport. Theoretically, hotel rent should be higher if it is closer to airport but our result shows the contrary

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

Table 7: Final Result Table

Variable	Coefficient	Std. Error	t-Statistic	Prob.
log(population)	-801.0531	2.13E+02	-3.768065	0.00
Star Rating [5]	11684.83	918.2034	12.72575	-
Star Rating [4]	5098.458	209.1757	24.37404	-
Star Rating [3]	3494.483	171.4874	20.37749	-
Star Rating [2]	2196.155	178.9652	12.27141	-
ISTOURISTDESTINATION	1195.803	218.628	5.469579	-
ISNEWYEAREVE	646.2088	219.3941	2.945425	0.00
HOTELCAPACITY	-13.21697	2.836197	-4.660104	-
HASSWIMMINGPOOL	2124.199	168.2217	12.62738	-
FREEWIFI	969.9183	248.9096	3.896669	0.00
FREEBREAKFAST	-485.7204	250.6377	-1.937938	0.05
AIRPORT	43.83494	1.06E+01	4.124019	-
C	10280.18	3071.344	3.347127	0.00

Room Rent = β_0 + β_1*Star Rating [1] + β_2*Is New Year Eve + β_3*Is Tourist Destination + β_4*City Population + β_5*Airport + β_6*Free Wifi + β_7*Free Breakfast + β_8*Hotel Capacity + β_9*Has Swimming Pool + β_{10}*Star Rating [2] + β_{11}*Star Rating [3] + β_{12}*Star Rating [3] + β_{13}*Star Rating [4]

REFERENCES

1. *Sharma, A., & Johri, A. (2012). FDI: An Instrument of Economic Growth & Development in Tourism Industry. International Journal of Scientific and Research Publications, 2(10), October 2012*
2. *https://in.hotels.com/ [Accessed on 18th Dec 2016]*
3. *Gujarati, Damodar, Basic Econometrics, Fifth Edition*
4. *Dougherty, Christopher, Introduction to Econometrics, Fourth Edition*
5. *Mir, Lateef, An Economic Evaluation of Indian Tourism Industry, International Journal of Scientific and Research Publications, Volume 4, Issue 12, December 2014*