**An Analysis Report**

**On**

**Hotel Room Pricing In the Indian Market**

**By**

**Sai Gopi Krishna Govindarajula**

**Under the supervision of**

**Prof. Sameer Mathur**

**Analysis of Hotel Room Pricing In the Indian Market**

**1. Introduction**

This report analyses the hotel rooms pricing In India based on various factors and how each of those factors affects the pricing. In this analysis we will consider the results from several graphs, histograms and plots drawn between the per day rent of various hotels versus various factors such as if the hotel has facilities like swimming pool, free Wi-Fi, free breakfast, and if the hotel capacity has any impact on the rent of the rooms etc.

The data used was collected from www.hotels.in in October 2016. The data collected is a field study which empirically investigates the pricing of hotel rooms located in 42 different cities of India during the time period of December 2016 to January 2016. We estimate a regression of hotel room prices in a mixed-model framework. Our model accounts for both fixed-effects and random-effects, controlled for unobserved heterogeneity. We estimate it using the Restricted Maximum Likelihood (REML) methodology. Our analysis reveals a significant “price of ratings, swimming pool and capacity” embedded in hotel room rent among hotels in different cities of India.

**2. An empirical field study of hotel pricing strategy in India**

**2.1 Hypotheses**

We study how the price of a room at a hotel is affected by external and internal factors. We can deduce that out of 18 factors, the 3 most influencing factors are Star Rating, Availability of swimming pool and Hotel capacity. We are taking these three factors on the basis of their strong correlation with room rent and we can see that in the correlation plots.

We will frame our hypothesis based on these three factors as well as with other binary functions which may affect the room rent of a hotel.

**H1:** *The average room rent in hotel which have swimming pools is more than that of hotels which don’t have swimming pools.*

**H2:** *The average Room Rent in hotels with high star rating is high as compared to ones which have less star rating.*

**H3:** *The average Room Rent in hotels with more capacity is lower than the Room Rent of hotels with less capacity.*

**H4:***The average Room Rent in hotels providing Free Breakfast is more than that which don’t provide breakfast.*

**H5:***The average Room Rent in hotels in metro cities is more than that of hotel in non metro cities.*

**2.2 Data**

The objective of this project is to analyze the pricing strategy of hotels in the Indian hotel industry. There are several factors that affect this pricing system. The target of this project is to identify the factors that matter the most and create models to estimate the pricing system. The dataset tracks hotel prices on 8 different dates at different hotels across different cities.

For this project, our dataset is based on hotels located in 42 Indian cities (Mumbai, Delhi, Bangalore, Chennai, Hyderabad, Ahmedabad, Kolkata, Surat, Pune, Jaipur, Thrissur, Lucknow, Kanpur, Amritsar, Indore, Kanyakumari, Agra, Madurai, Goa, Rajkot, Varanasi, Srinagar, Jodhpur, Chandigarh, Thiruvathipuram, Guwahati, Mysore, Bhubaneswar, Kochi, Mangalore, Udaipur, Pondicherry, Haridwar, Puri, Shimla, Panchkula, Darjeeling, Rishikesh, Gangtok, Ooty, Jaisalmer, Bodh Gaya, Nainital, Munnar, Manali) India.

We collected data from the well-known website [www.hotels.in](http://www.hotels.in) that aggregates the hotel prices on 8 different dates at different hotels across different cities. There is a good chance that lot of factors govern the rent of these hotel rooms. Any meaningful empirical analysis will need to control for factors. For example, factors such as whether the hotel is rated as a five star hotel, how many rooms does it have, has swimming pool are all likely to influence hotel prices.

**Price:** We collected data from 18 Dec 2016 to 08 Jan 2017. We used RoomRent to denote the average price of a room at a hotel. We measured Room Rent, rent for the cheapest room, double occupancy, in Indian Rupees. Some hotels have more than one type of double occupancy room. For simplicity, we picked the cheapest room with double occupancy at hotel k in city j.

**Star Rating**:

In India, the Ministry of Tourism has formulated a scheme for classification of operational hotels using a “Star” rating. Hotels are rated as either 5 Star, 4 Star, 3 Star, 2 Star or 1 Star. Accordingly, we classified the hotels in our dataset using their star rating. The reason for doing this is that the star rating of a hotel has a direct, strongly positive correlation with the price of its hotel rooms. Therefore, it is important to control for price variation because of the star rating. We used the variable StarRating to denote the star rating of hotel k in city j.

**Hotel Capacity:**

We recorded the total number of rooms in hotel k in city j as HotelCapacity. Ultimately, the number of rooms in a hotel denotes the available supply and it is expected that this will keenly influence the price that a hotel will set. Accordingly, we used HotelCapacity as a control variable to account for the possibility that the room price set by a hotel may depend upon the supply of available rooms.

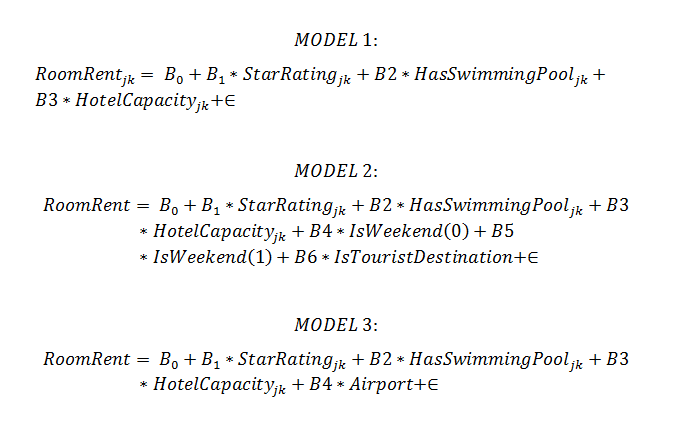
**Swimming Pool:**

The amenities and facilities provided within a hotel can also potentially influence the price of a room. The greater the amenities, the higher should be the price of the hotel room. To partially control for such factors, we recorded whether a hotel had a Swimming Pool or not. We used HasSwimmingPool to denote the presence or absence of a Swimming Pool at hotel k in city j.

**2.3 Model**

We analyzed the research question using one model.

**Model:** We established the effect of Star Rating, Hotel Capacity and availability of Swimming Pool on the price of a room in a hotel with the simplest model we could come up with. We regressed the room rent on the variables Star Rating, Hotel Capacity and whether hotel had a swimming pool, in our second model the previous three variables remained and we added IsTouristDestination and IsWeekend as factors and lastly we added Airport distance from the hotel in the basic three variable of the beginning to propose a better model, they are as follows.



We estimated three different Models, described above using linear least squares.

The benefit of having the three regressors outlined in Model was that it helped us rule out some alternate explanations for the variation in hotel room rent. For example, it is well-known that five-star hotels are more expensive than four-star hotels. Including the star rating as a regressor, permitted us to investigate the effect of other variables on hotel room rent, after controlling for price variation due to the star rating. We expected to find the coefficient for StarRating to be positive (B1>0). Similarly, having a dummy variable has Swimming Pool or not for each hotel, permitted us to control effect of availability of swimming pool on rent of hotel rooms and the same way about Hotel capacity, whether the place is a tourist destination, what is the distance of hotel from the airport, the sold out of all rooms in hotel depends on weekend or not, etc.

**2.4 Results**

**Model:** The analysis of Model also yielded statistical support for our hypotheses H1s. Recall that Model includes three to six independent variables, as shown in equations above. We found that the average room prices with higher ratings and having swimming pool were higher than the prices with low ratings and no swimming pool. This regression analysis yielded B1 >0, B2 >0, with p <0.05, as shown in Table 2. As expected, we additionally observed a negative relationship between the average hotel room prices and the hotel capacity, B3>0, with p < 0.05. But, we found that the Model that we have taken doesn’t have very good R², so there may exist models better than the model we have taken with other variables, in explaining the relationship between hotel pricing strategies.

The coefficients and linear model of the above three models mentioned in 2.3are as follows:-

# Model1: Hotel Rent = b0 + b1\*StarRating + b2\*HasSwimmingPool+ b3\*HotelCapacity

# b0 = -6896.154, b1 = 3597.322, b2=2528.885, b3= -15.558

#Model2: Rent= b0 + b1\*StarRating + b2\*HasSwimmingPool+ b3\*HotelCapacity +b4\*IsWeekend(0) + b5\*IsTouristDestination

#b0 = -1(assumption), b1 = 1258.9558 , b2=3670.2511, b3= -6.1769 , b4= -509.6479, b5=1053.0394

# Model3: Rent = b0 + b1\*StarRating + b2\*HasSwimmingPool+ b3\*HotelCapacity +b4\*Airport

# b0 = -7288.04830, b1 = 3522.99002 , b2=2708.40013, b3= -14.77562, b4= 25.34377

The coefficients of model 1 implies:

* When Star rating increases by 1 unit Hotel Room rent increases by Rs.3597.32
* When availability of swimming pool changes then Hotel room rent increases by Rs 2528.88
* When Hotel capacity changes by 1 unit then Hotel room rent decreases by Rs 15.55.

The coefficients of model 2 implies:

* When Star rating increases by 1 unit Hotel Room rent increases by Rs.1258.95
* When availability of swimming pool changes then Hotel room rent increases by Rs 3670.25
* When Hotel capacity changes by 1 unit then Hotel room rent decreases by Rs 6.17
* Depending on whether there is a weekend or not, the room rent varies by Rs. 509.64
* Is hotel is near a tourist destination then room rent increases by

Rs 1053.03.

The coefficients of model 3 implies:

* When Star rating increases by 1 unit Hotel Room rent increases by Rs 3522.99
* When availability of swimming pool changes then Hotel room rent increases by Rs 2708.40
* When Hotel capacity changes by 1 unit then Hotel room rent decreases by Rs 14.77
* When distance of airport from any hotel increases by 1 unit then Hotel room rent increases by Rs 25.34

**3. Conclusion**

This report was motivated by the need for research that could improve our understanding of how different external and internal factors influences the pricing strategies in the hotel industry of India. The unique contribution of this paper is that we investigated the price premium charged by hotels according to the facilities they provide and also where it is situated.

This research has some important managerial implications. We found that not only supply affects prices but also there many other factors which can influence our pricing strategy. When consumer sees good ratings and reviews about hotel and gets better amenities, it prompts an increase in quality perceptions, purchase intentions and willingness-to-pay.

**Table 1:** Summary Statistics of Hotels pricing strategy study

Room Rent

Min. 1st Qu. Median Mean 3rd Qu. Max.

299 2436 4000 5474 6299 322500

Star Rating

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.000 3.000 3.000 3.459 4.000 5.000

Swimming Pool

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.0000 0.0000 0.0000 0.3558 1.0000 1.0000

Hotel Capacity

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.00 16.00 34.00 62.51 75.00 600.00

Is a Tourist Destination

Min. 1st Qu. Median Mean 3rd Qu. Max.

1. 0.0000 0.0000 0.2842 1.0000 1.0000

Is a Weekend when all rooms sold out

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.0000 0.0000 1.0000 0.6972 1.0000 1.0000

Airport distance from the hotel

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.20 8.40 15.00 21.16 24.00 124.00

**Table 2:** Regression Analysis in the Hotels pricing strategy study

**Model1**

Call:

lm(formula = RoomRent ~ StarRating + HasSwimmingPool + HotelCapacity,

data = hotelData.df)

Residuals:

Min 1Q Median 3Q Max

-10804 -2295 -946 1002 310110

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -6896.154 340.549 -20.25 <2e-16 \*\*\*

StarRating 3597.322 111.670 32.21 <2e-16 \*\*\*

HasSwimmingPool 2528.885 157.894 16.02 <2e-16 \*\*\*

HotelCapacity -15.558 1.006 -15.47 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6710 on 13228 degrees of freedom

Multiple R-squared: 0.1628, Adjusted R-squared: 0.1626

F-statistic: 857.5 on 3 and 13228 DF, p-value: < 2.2e-16

P-values of individual independent variables are also < 0.05, So we reject all the null hypotheses and hence, Star Rating, availability of swimming pool and Hotel Capacity all significantly affects Room Rent.

R2 is very small so it may not be good model and some other factors which are not included might affect Room rent.

**Model 2:**

Call:

lm(formula = RoomRent ~ StarRating + HasSwimmingPool + HotelCapacity +

IsWeekend + IsTouristDestination - 1, data = hotelData.df)

Residuals:

Min 1Q Median 3Q Max

-8326 -2517 -1212 463 312480

Coefficients:

Estimate Std. Error t value Pr(>|t|)

StarRating 1258.9558 44.4985 28.292 < 2e-16 \*\*\*

HasSwimmingPool 3670.2511 148.8411 24.659 < 2e-16 \*\*\*

HotelCapacity -6.1769 0.9658 -6.396 1.65e-10 \*\*\*

IsWeekend -509.6479 119.1618 -4.277 1.91e-05 \*\*\*

IsTouristDestination 1053.0394 124.7325 8.442 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6792 on 13227 degrees of freedom

Multiple R-squared: 0.4493, Adjusted R-squared: 0.4491

F-statistic: 2159 on 5 and 13227 DF, p-value: < 2.2e-16

**Model 3:**

Call:

lm(formula = RoomRent ~ StarRating + HasSwimmingPool + HotelCapacity +

Airport, data = hotelData.df)

Residuals:

Min 1Q Median 3Q Max

-10785 -2265 -876 982 310437

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -7288.048 341.691 -21.329 <2e-16 \*\*\*

StarRating 3522.990 111.531 31.588 <2e-16 \*\*\*

HasSwimmingPool 2708.400 158.397 17.099 <2e-16 \*\*\*

HotelCapacity -14.776 1.006 -14.695 <2e-16 \*\*\*

Airport 25.344 2.590 9.786 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6687 on 13227 degrees of freedom

Multiple R-squared: 0.1688, Adjusted R-squared: 0.1686

F-statistic: 671.7 on 4 and 13227 DF, p-value: < 2.2e-16