Analysis of Hotel Room Pricing In the Indian Market

### Name: Pratyusha Maiti

### Email: [pratyusha.maiti@gmail.com](mailto:pratyusha.maiti@gmail.com)

### College: Jadavpur University

## A Study of Hotel Room Pricing Strategy in the Indian Market

## **1. Introduction**

The analysis is on the pricing of hotel rooms in India based on certain internal and external factors under study.

A number of factors are studied and their relationship with room rent is visualised by means of graphs, histograms and plots. We take up special complementary features hotels often boast about, such as swimming pools and free WiFi and attempt to gauge how rents vary with such luxuries. We also take up the possibility of room rents varying with special occasions or dates, or with the city.

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. This accounts for variety and a fairly general conclusion can be drawn from the analysis.

After studying myriad factors and combinations, we find a high dependence of pricing of hotel rooms on the hotel star rating, its capacity and whether or not the city is a tourist destination. Based on our findings, regression models are prepared to represent the data as well as possible.

## **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. From Boruta test we find that factors such as Weekend, Date and New Year Eve have least effect on the room rent hence can be rejected. Out of 18 factors, on further correlation tests, the 3 most influencing factors are star rating, whether or not the city is a tourist destination and hotel capacity. We are taking these three factors on the basis of a strong correlation with room rent.

Subsequently, we frame certain hypothesis based on these three factors as well as with other factors which may affect the room rent of a hotel, thinking intuitively.

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

**H2:** *The average Room Rent of hotels located closer to the airport are higher than those that are further away.*

**H3:** *The average Room Rent in hotels with high star rating is high as compared to one which has less star rating.*

**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.*

**H6:** *The average Room Rent in hotels with more capacity is lower than the Room Rent with low hotel capacity.*

## **2.2 Data**

The purpose of this project is to analyse the pricing strategy of hotels in the Indian hotel industry. Many factors drive hotel room prices. We wish to identify the factors that matter the most and build a corresponding regression model. The dataset tracks hotel prices on 8 different dates at different hotels spread across 42 cities in India.

The 42 Indian cities under consideration are 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. The data is collected from the well-known website[www.hotels.in](http://www.hotels.in).

A glance at the data set shows that many factors may govern the renting of hotel rooms. For a meaningful empirical analysis, the factors need to be managed. For example, factors such as whether the hotel is rated as a five star hotel, how many rooms does it have, distance from airport, whether it has a swimming pool or the city being a tourist destination are all likely to influence hotel prices, but we need to find out what best describes the pricing strategy without need for including excess and redundant terms.

**Price:** Collected data ranges for dates 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.

**Star Rating**:

The “Star” rating is a scheme for classification of operational hotels formulated by the Ministry of Tourism in India. Hotels are rated as either 5 Star, 4 Star, 3 Star, 2 Star or 1 Star in accordance with various customer satisfaction factors. Accordingly, we classified the hotels in our dataset using their star rating, as variable StarRating. If we think about it, the star rating is a strong pointer to a hotel’s maintenance, facilities and functionality. Hence, the star rating of a hotel has a direct, strongly positive correlation with the price of its hotel rooms.

**Hotel Capacity:**

The total number of rooms in a hotel, HotelCapacity 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. This again gives room to think that this demand would increase on special days.

**Distance from airport:**

The distance of a hotel from airport is often a factor people add to their checklist when hotel-hunting. Hotels closer to airports guarantee a certain degree of convenience for travelling. Strategically, hiked prices correspond to ease of travel and thus is a possible important factor. We use Airport to denote the distance between the hotel and closest major airport.

**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.

## **2.3 Model**

We analyzed the research question using a basic primary model with three major terms.

**Model :** We established the effect of Star Rating, Hotel Capacity and distance from closest major airport on the price of a room in a hotel with a simple model. We regressed the room rent on the variables Star Rating, Hotel Capacity and distance from airport, in our primary model. Subsequently, more number of terms were added to improve the model, based on Residual Standard Error and R-squared terms.

We estimated six different Models, in accordance with above using linear least squares.

The benefit of having multiple 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 like 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, add to the control we have over refining a model that closely describes the pricing strategy of hotels.

**2.4 Results**

**Model:** From the six models that we analysed, four models are important as they show statistically how the factors improve the overall model. The remaining two models were constructed to show how certain factors that seem important, do not contribute as much when statistically evaluated. We will discuss the former four models in detail.

The four major models under study are:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model 1 | *salary = b0 + b1\*StarRating + b2\*IsTouristDestination+ b3\*HotelCapacity*   |  |  | | --- | --- | | b0 = | -1 (estimated) | | b1 = | 4.348e+03 | | b2 = | 2.071e+03 | | b3 = | -1.107e+01 | |
| Model 2 | *salary = b0 + b1\*StarRating + b2\*IsTouristDestination + b3\*HasSwimmingPool + b4\*HotelCapacity*   |  |  | | --- | --- | | b0 = | -8351.683 | | b1 = | 3635.967 | | b2 = | 1877.481 | | b3 = | 2285.337 | | b4 = | -13.965 | |
| Model 3 | *salary = b0 + b1\*StarRating + b2\*Airport + b3\*HasSwimmingPool + b4\*HotelCapacity + b5\*HotelPincode + b6\*IsTouristDestination + b7\*CityRank + b8\*Population + b9\*FreeWifi + b10\*IsMetroCity*   |  |  | | --- | --- | | b0 = | -8.374e+03 | | b1 = | 3.597e+03 | | b2 = | 1.042e+01 | | b3 = | 2.157e+03 | | b4 = | -1.041e+01 | | b5 = | -1.858e-04 | | b6 = | 1.884e+03 | | b7 = | 6.582e-01 | | b8 = | -1.269e-04 | | b9 = | 5.987e+02 | | b10 = | -6.159e+02 | |
| Model 4 | *salary = b0 + b1\*StarRating + b2\*StarRating:HasSwimmingPool + b3\*StarRating:FreeWifi + b4\*HasSwimmingPool:FreeWifi + b5\*StarRating:HasSwimmingPool:FreeWifi + b6\*Airport-1 + b7\*HasSwimmingPool + b8\*HotelCapacity + b9\*HotelCapacity:StarRating + b10\*IsTouristDestination + b11\*CityRank + b12\*Population + b13\*FreeWifi + b14\*IsMetroCity*   |  |  | | --- | --- | | b0 = | 0 | | b1 = | 1.004e+03 | | b2 = | 9.320e+00 | | b3 = | -7.218e+03 | | b4 = | 4.058e+00 | | b5 = | 1.860e+03 | | b6 = | 8.874e+00 | | b7 = | -9.794e-05 | | b8 = | -4.553e+03 | | b9 = | -7.618e+02 | | b10 = | 2.548e+03 | | b11 = | 1.445e+03 | | b12 = | -7.759e+02 | | b13 = | -3.883e+00 | | b14 = | 3.422e+02 | |

**Table 1: Summary statistics of hotel pricing strategy analysis**

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 |

Airport

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| 0.20 | 8.40 | 15.00 | 21.16 | 24.00 | 124.00 |

Hotel Capacity

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| 0.00 | 16.00 | 34.00 | 62.51 | 75.00 | 600.00 |

**Table 2: Regression models of hotel pricing strategy analysis**

**Model 1:**

Call:

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

data = cities.df)

Residuals:

Min 1Q Median 3Q Max

-10763 -2466 -863 964 309881

Coefficients:

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

(Intercept) -1.032e+04 3.274e+02 -31.52 <2e-16 \*\*\*

StarRating 4.348e+03 1.001e+02 43.42 <2e-16 \*\*\*

IsTouristDestination 2.071e+03 1.275e+02 16.24 <2e-16 \*\*\*

HotelCapacity -1.107e+01 9.913e-01 -11.16 <2e-16 \*\*\*

---

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

Residual standard error: 6709 on 13228 degrees of freedom

Multiple R-squared: 0.1633, Adjusted R-squared: 0.1631

F-statistic: 860.3 on 3 and 13228 DF, p-value: < 2.2e-16

**Model 2:**

Call:

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

HotelCapacity, data = cities.df)

Residuals:

Min 1Q Median 3Q Max

-11206 -2380 -716 1075 309612

Coefficients:

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

(Intercept) -8351.683 351.902 -23.73 <2e-16 \*\*\*

StarRating 3635.967 110.797 32.82 <2e-16 \*\*\*

IsTouristDestination 1877.481 127.240 14.76 <2e-16 \*\*\*

HasSwimmingPool 2285.337 157.484 14.51 <2e-16 \*\*\*

HotelCapacity -13.965 1.004 -13.91 <2e-16 \*\*\*

---

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

Residual standard error: 6656 on 13227 degrees of freedom

Multiple R-squared: 0.1764, Adjusted R-squared: 0.1761

F-statistic: 708.1 on 4 and 13227 DF, p-value: < 2.2e-16

**Model 3:**

Call:

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

HotelCapacity + HotelPincode + IsTouristDestination + CityRank +

Population + FreeWifi + IsMetroCity)

Residuals:

Min 1Q Median 3Q Max

-11475 -2376 -706 1036 309427

Coefficients:

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

(Intercept) -8.374e+03 4.727e+02 -17.716 < 2e-16 \*\*\*

StarRating 3.597e+03 1.108e+02 32.464 < 2e-16 \*\*\*

Airport 1.042e+01 3.337e+00 3.122 0.001801 \*\*

HasSwimmingPool 2.157e+03 1.618e+02 13.329 < 2e-16 \*\*\*

HotelCapacity -1.041e+01 1.030e+00 -10.111 < 2e-16 \*\*\*

HotelPincode -1.858e-04 2.474e-04 -0.751 0.452475

IsTouristDestination 1.884e+03 1.497e+02 12.587 < 2e-16 \*\*\*

CityRank 6.582e-01 1.060e+01 0.062 0.950496

Population -1.269e-04 3.784e-05 -3.354 0.000799 \*\*\*

FreeWifi 5.987e+02 2.219e+02 2.697 0.006995 \*\*

IsMetroCity -6.159e+02 2.194e+02 -2.807 0.005001 \*\*

---

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

Residual standard error: 6606 on 13221 degrees of freedom

Multiple R-squared: 0.189, Adjusted R-squared: 0.1884

F-statistic: 308.1 on 10 and 13221 DF, p-value: < 2.2e-16

**Model 4:**

Call:

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

StarRating:FreeWifi + HasSwimmingPool:FreeWifi + StarRating:HasSwimmingPool:FreeWifi +

Airport - 1 + HasSwimmingPool + HotelCapacity + HotelCapacity:StarRating +

IsTouristDestination + CityRank + Population + FreeWifi +

IsMetroCity)

Residuals:

Min 1Q Median 3Q Max

-13499 -2222 -635 1034 307852

Coefficients:

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

StarRating 1.004e+03 1.204e+02 8.344 < 2e-16 \*\*\*

Airport 9.320e+00 3.154e+00 2.955 0.003134 \*\*

HasSwimmingPool -7.218e+03 1.908e+03 -3.783 0.000156 \*\*\*

HotelCapacity 4.058e+00 5.839e+00 0.695 0.487047

IsTouristDestination 1.860e+03 1.465e+02 12.695 < 2e-16 \*\*\*

CityRank 8.874e+00 1.024e+01 0.867 0.386086

Population -9.794e-05 3.546e-05 -2.762 0.005756 \*\*

FreeWifi -4.553e+03 5.370e+02 -8.479 < 2e-16 \*\*\*

IsMetroCity -7.618e+02 2.147e+02 -3.549 0.000388 \*\*\*

StarRating:HasSwimmingPool 2.548e+03 4.818e+02 5.289 1.25e-07 \*\*\*

StarRating:FreeWifi 1.445e+03 1.926e+02 7.501 6.73e-14 \*\*\*

HasSwimmingPool:FreeWifi -7.759e+02 1.999e+03 -0.388 0.697925

StarRating:HotelCapacity -3.883e+00 1.260e+00 -3.081 0.002067 \*\*

StarRating:HasSwimmingPool:FreeWifi 3.422e+02 5.076e+02 0.674 0.500279

---

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

Residual standard error: 6555 on 13218 degrees of freedom

Multiple R-squared: 0.4873, Adjusted R-squared: 0.4868

F-statistic: 897.5 on 14 and 13218 DF, p-value: < 2.2e-16

But, these models that we have taken do not 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.

## **3. Conclusion**

This paper serves to show how different external and internal factors influence the pricing strategies of rooms in the hotels in India. Keeping in mind a managerial approach to the conclusions, we will now discuss our findings on pricing of rooms and which factors affect it most.

From the above results, it is clear that the last model has the best fit among all the models. However, we see, just a few simple terms do not give this model, instead there are terms dependent on each other which gives the best value.

We find that the single most important factor is the star rating of the hotels. This is obvious, considering that the star rating is what attracts customers to try a hotel, in turn increasing the popularity and demand for rooms.

Another important factor is the distance to airport, as mentioned earlier. It is also seen that tourist attraction areas tend to have higher priced hotels. Again, this is a simple matter of demand.

More complex terms come into picture like city rank and population. A positive dependence on city rank is clear since higher ranking cities are expected to have better hotel rooms.

From the managerial point of view, in order to make your rooms worth a higher price point, the hotel focus on:

* Location: The peak lies at high rank cities of tourist destination, and the hotel needs to be situated near an airport.
* Hotel Interior: Must have good capacity, and added facilities that justify the price point. However, unnecessary features just for hiking price is a bad move.
* Maintain customer satisfaction: This is intuitively the most important factor, as word-of-mouth and ratings are best tools in drawing more customers.



[Hotel Rambagh Palace](https://taj.tajhotels.com/en-in/taj-rambagh-palace-jaipur/), Jaipur: Highest priced rooms at 3 lakhs per day.