**A Study of Hotel Room Pricing in Indian cities**

**1. Introduction**

Hotel room pricing is a delicate issue for any issue because even a slight difference in pricing with its competitors can lead to heavy loses. And even setting the price too low can hamper their business. Hotel business flourish in two types of area i.e. tourist destinations and metro cities. In this regard, suitably pricing hotel rooms and related services for customers becomes a crucial issue. This is because a hotel’s price reflects an assessment of the value that customers see and their willingness-to-pay for the hotel’s rooms and services.

This paper addresses the following issues concerning the “price of rooms” with respect to the hotel industry. The pivotal issue that concerns the industry is the hotel pricing strategy. In this paper, we investigate what internal and external factors determine the room of a price? We evaluate whether hotels extract a “price of a holiday” when the traffic is high, compared to the usual settings?

The second issue concerns consumer perceptions and willingness-to-pay for hotel rooms in unconventional settings versus conventional settings.

Our field of study will be limited to a total of 42 cities in India. All the major metros, heritage destinations and places with booming natural sceneries have been covered here. So, we do have a large sample space to find something conclusive. We have a dynamic range of hotels, starting from 1 star hotel to 5 star hotels, spread all over the country. 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.

**2. An empirical field study of hotel prices in India**

Many local people have started taking advantage of the major tourist attractions that they have in their town. They have started converting their homes into a seasonal hotel, earning hefty profits due to lack of competition by big players. Their pricing strategy is also governed by the influx of customers they are receiving.

**2.1 Hypothesis**

We first study the relationship between the hotel room rent for a hotel with or without a swimming pool. It might seem like an absurd regressor variable to take while comparing hotel prices but it is important because a higher star rating hotel is bound to have a swimming pool. This observation is based purely on intuition.

**H1:** *The prices of hotel rooms at hotels with a swimming pool are higher than the prices at hotels without one.*

**H2:** *The prices of hotel rooms at hotels is higher on a weekend than on a regular weekday.*

**2.2 Data**

We have taken data from 42 different cities. We have tourist places with large and small population and also metro cities. The cities have hotels located in traditional, tourist sites, as well as conventional modern hotels. We collected data from the website www.hotels.in that aggregates hotel, room prices and features and the city population, type and rankings.

It is indeed probable that many factors outside the listed internal and external factors govern the prices of hotel rooms. Any meaningful empirical analysis will need to control for factors. For example, factors such as the city the hotel is located in, whether the hotel is rated as a five star hotel, how many rooms does it have, how far is it from the airport and train station are all likely to influence hotel prices. Some possible factors not listed here are the accessibility of the location, the presence of an in-house restaurant and the online pre-booking facilities provided by the hotel.

**City:** It is likely that the city in which a hotel is located in will strongly influence the hotel room prices. We collected data from forty two Indian cities. Specifically, we used a dummy variable *CityRank* to index the cities according to their population. We have not used any more variables for particular hotels. We have been given their name and the description of the room.

**Type:** We have used two variables to bifurcate the cities. The first one is *IsMetroCity* where 0 defines a non-metro city while 1 is assigned to a metro city. Similarly the second important divide that has been created is based on the tourist significance of the destination. Variable *IsTouristDestination* determines whether a city has been more linked to tourism or not based on similar assignment of 0 and 1.

**Date:** The time when you visit a place, can also cause fluctuation in the prices of hotel rooms because hotel availability in less in holiday season and hotel managers tend to increase the price of hotel rooms, seeing the higher demand. Thus, we have two more variables to account for this change. *IsWeekend* and *IsNewYearEve* are the two variables.

**Price:** The data was collected on October 2016. We have used *RoomRent* to determine the price of the cheapest double occupancy hotel room. This has been done for simplicity in comparison.

**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 have used the variable *StarRating* for this purpose.

**Rooms:** We recorded the total number or rooms in hotels and will access this number using the variable *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. **Distance from the Airport:** It is possible that hotels located close to the airport are able to charge a price premium for the greater convenience and easy access. In order to control for this alternate explanation, we recorded the distance between a given hotel and the closest airport and railway station. We have used the variable *Airport* which denotes the distance between a hotel and the nearest airport to the hotel.

**Amenities:** 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 control the prices of a hotel room, we have introduced three dichotomous variables which are *FreeWifi, FreeBreakfast* and *HasSwimmingPool.* The first two amenities can usually be found in smaller hotels with lesser ratings too. But having a swimming pool inside the premises of a hotel can only be pulled off by highly rated hotels. This could have a significant impact on the final pricing of the hotel room.

Table 1 shows the summary statistics.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | vars | n | mean | sd | median |
| Hotel Number | 1 | 13232 | 6616.5 | 3819.893716 | 6616.5 |
| CityName\* | 2 | 13232 | 18.06650544 | 11.71739595 | 16 |
| Population | 3 | 13232 | 4416836.872 | 4258385.997 | 3046163 |
| CityRank | 4 | 13232 | 14.8337364 | 13.51245788 | 9 |
| IsMetroCity | 5 | 13232 | 0.284159613 | 0.451030267 | 0 |
| IsTouristDestination | 6 | 13232 | 0.697173519 | 0.459498161 | 1 |
| IsWeekend | 7 | 13232 | 0.622808343 | 0.484701832 | 1 |
| IsNewYearEve | 8 | 13232 | 0.124395405 | 0.330044574 | 0 |
| Date\* | 9 | 13232 | 14.25816203 | 2.81790756 | 14 |
| HotelName\* | 10 | 13232 | 841.1877267 | 488.1569964 | 827 |
| RoomRent | 11 | 13232 | 5473.991838 | 7333.116787 | 4000 |
| StarRating | 12 | 13232 | 3.45893289 | 0.756232469 | 3 |
| Airport | 13 | 13232 | 21.15874395 | 22.75990616 | 15 |
| HotelAddress\* | 14 | 13232 | 1202.530985 | 582.1659918 | 1261 |
| HotelPincode | 15 | 13232 | 397430.2646 | 259837.4973 | 395003 |
| HotelDescription\* | 16 | 13224 | 581.3383999 | 363.2607483 | 567 |
| FreeWifi | 17 | 13232 | 0.925861548 | 0.262005973 | 1 |
| FreeBreakfast | 18 | 13232 | 0.649108222 | 0.477267171 | 1 |
| HotelCapacity | 19 | 13232 | 62.51163845 | 76.6633426 | 34 |
| HasSwimmingPool | 20 | 13232 | 0.355804111 | 0.478774341 | 0 |

**3. Models**

We analyzed the research question using a multiple variable linear model.

**Model 1:** We first established the effect of having a swimming pool on the price of a room in a hotel with the simplest model we could come up with. We regressed the price of the hotel as follows.

We estimated Model 1, using linear least squares. If there was a price of having a swimming pool, we expected to find >0 in support of hypotheses H1.

**Model 2:** We first established the effect of the date being a weekend or not on the price of a room in a hotel with the simplest model we could come up with. We regressed the price of the hotel as follows.

We estimated Model 1, using linear least squares. If there was a price of the date of hotel reservation, we expected to find >0 in support of hypotheses H2.

**Model 3:** Next, as a robustness check, we defined a detailed model accounting for eleven additional independent variables, which may also influence the variation in hotel prices. Our revised regression model was as follows.

+

We estimated Model 3, described above using linear least squares. Once again, if there was indeed a “price of swimming pool” in the hotel market, we expected to find >0 in support of hypotheses H1.

Another benefit of having the eleven additional regressors outlined in Model 2 was that it helped us rule out some alternate explanations for the variation in hotel prices. Similarly, having a dummy variable for each city, permitted us to control for city-wide variation in prices of hotel rooms, potentially arising out of differences in real-estate prices and other expenses across cities.

**4.1 Results**

**Model 1:** We found empirical support for H1. The average room prices at hotels with swimming pool were higher than the prices at hotels without one. The regression analysis using Ordinary Least Squares yielded >0, with p < 0.05.

**Model 2:** We could not find empirical support for H2. We fail to reject the hypothesis that the average prices of the hotels in the weekend is same as that on the weekday. The regression analysis using Ordinary Least Squares yielded <0, with p > 0.05, which was well outside the confidence interval.

**Model 3:** The analysis of Model 2 also yielded statistical support for our hypothesis H1 and not in favor of H2. Extending the above two models, we again found that the average room prices at hotels with a swimming pool were higher than the prices at hotels without one and hotel pricing didn’t change significantly over the weekend. This regression analysis also yielded >0, with p <0.05 and <0, with p >0.05. As expected, we additionally observed a positive relationship between the average hotel room prices and the hotel star ratings, >0, with p < 0.001. Model 3 fit the data better than Model 1 and Model 2. Overall, we found Model 3 to be better in explaining the relationship between hotel pricing and the other factors involved.

**4.2 Interpretation**

Using the predictor equation of Model 3, we can draw the following inferences:

1. Since negative hotel room rent is impossible, we cannot give a physical interpretation to the intercept. It is merely an adjustment constant.
2. The regression coefficients (1334.7866) of *StarRating* is significantly different from zero, (p < 0.001) and indicates that there is an expected increase of 1334.7866 amount of room rent for every one star rating increase of the hotel. Similar deductions can be made for other variables too where p < 0.001.
3. The multiple R-squared (0.3902) is best estimated by Model 3. It means model accounts for 39.02% of the variance in room rent.
4. The residual standard error is the average error in predicting room rent from the Model 3.

**4.3 Conclusion**

This paper was motivated by the need to understand the pricing strategy of hotels in several different cities in India. Two important observations that we found from the investigation was that the pricing was directly dependent on the star rating of the hotel but it was not dependent on whether the date of reservation is a weekend or a weekday.

This research has some important managerial implications. We found that customers value are willing to shell out more money in Tourist Destinations and the Star Rating of the hotel is important for determining the room rent of the hotel. Now that we have seen the data and made relevant inferences, a hotel manager could set the price of the hotel room accordingly to earn a certain margin of profit and crowd.

Table 2 shows the summary of the fitted model 3.

|  |
| --- |
| Coefficients Estimate Std. Error t value Pr(>|t|) |
| (Intercept) -1766.9942 130.7291 -13.516 < 2e-16 |
| StarRating 1334.7866 35.3549 37.754 < 2e-16 |
| CityRank 29.9464 1.9704 15.198 < 2e-16 |
| HasSwimmingPool 929.1030 48.8431 19.022 < 2e-16 |
| IsWeekend -49.0553 37.5640 -1.306 0.192 |
| IsMetroCity 228.7278 52.1719 4.384 1.17e-05 |
| IsTouristDestination 447.6307 44.1442 10.140 < 2e-16 |
| IsNewYearEve 265.9258 55.9234 4.755 2.00e-06 |
| Airport 14.8541 0.9042 16.429 < 2e-16 |
| FreeBreakfast 419.2887 37.4760 11.188 < 2e-16 |
| FreeWifi -416.6564 67.3995 -6.182 6.53e-10 |
| HotelCapacity 4.1531 0.3326 12.487 < 2e-16 |