**A Study of Hotel Pricings in India**

1. **Introduction**

Tourism has now become a significant industry in India. It is a sun rise industry, an employment generator, a significant source of foreign exchange for the country. Tourism in India is the third largest foreign exchange earner of the country. The booming tourism industry has had an interleaving effect on the hospitality sector with an increase in the occupancy ratios and average room rates.

As per World Travel and Tourism Council (WTTC), India is one of the favourite tourist destinations from the year 2009 and will continue to be one of the favourite till 2018. Further, the Travel and Tourism Competitiveness Report by World Economic Forum, has ranked India at the sixth place in tourism and hospitality. The tourism and hospitality sector is among the top 10 sectors in India to attract the highest Foreign Direct Investment (FDI). According to the data released by Department of Industrial Policy and Promotion (DIPP), the hotel and tourism sector attracted around US$ 9.2 billion of FDI between April 2000 and March 2016.

The hotel industry in India thrives largely due to 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.

This paper addresses the issues concerning with “Hotel Pricings in India”, therefore we are going to put forth a few parameters that have an influence on the room rents pertaining to various cities in India. **We investigate the significant factors that help to predict the room rent of a particular hotel.** We consider a few factors like, distance from airport, free Wi-Fi, free breakfast, has swimming pool and many such parameters to denote the effect of those parameters on the room rent.

1. **Overview of the Study**

Our study is about the various pricings of different hotels in each of the 42 cities given in the data set. Our study tells us about the factors that help to show relevance to the pricings offered. A few hotels are more expensive than a few others so this study helps us to elaborate as to why a pricing difference comes between hotels that all serve the same purpose. We do have a wide variety of subsections to pick out from, based on metropolitan cities, cities that are tourist spots etc… which help us to analyse the overall condition of the city on its own. Secondly, it’s about the hotels themselves, that is, their distance to the airport, number of rooms available, type of rooms, whether they have free Wi-Fi, or free breakfast of swimming pool. Also, there are other situations when there is a hike in hotel prices that is during a weekend or Christmas eve or new year’s eve, in such scenarios also we will be analysing its effect on the hotel room rents. We compare the prices of different hotels across different cities all over India ([www.hotels.in](http://www.hotels.in)). We empirically study how the above mentioned parameters actually affect the prices of the rooms. We do notice a few trends in the data and that the rooms are all each described by having two rooms or the star rating or the number of restaurants that they possess. A number of these factors are taken into consideration. For instance, most of the room rents are high for those hotels that are placed closer to the airport. The highest of the prices ranges go for cities like Udaipur, Jodhpur , Srinagar which are heavy destination spots and least prices range to places like Kanpur, Bhubaneshwar etc.. which are neither metropolitan cities nor tourist spots. Also, other metropolitan cities have higher room rents like Chennai, Mumbai, Delhi etc.. Therefore, our analysis helps us to ponder over the very many factors that may contribute to a higher room rent than other hotels in India.

1. **An empirical field study of hotel prices all over India**
   1. **Hypothesis**

The investigation to be made here is to find out the various factors that are involved in the price changes offered at various parts of the country. Here, we take into consideration 42 cities that have fallen in the category of renowned cities in India. We have considered various external factors such as whether it’s a metropolitan city, whether it’s a tourist spot and the rank of the city. Therefore, these factors contribute to the pricings. There has been a general trend observed that shows that hotels cost higher at tourist spots and metropolitan cities than those that do not possess such qualities. Secondly, the cities ranked higher are usually posed with higher rents than usual. Other factors that are being considered here is the availability of a few features in the hotels which we will be elaborating later.

*H1: Hotels placed in metropolitan cities, tourist spots (both also) as well as those that are ranked higher have higher room rents than those that do not possess the above.*

With such a huge data set with both external and internal factors that affect the data set, we have a number of other features of the hotel that do contribute to the increase in the room rents than the usual prices. In this data set we have a total of 11 internal factors out of which those with higher rating, closer to the airport, hotel capacity and having a swmming pool are strongly responsible for room rents. Although, free wifi and free breakfast have a p value a little bit more than 0.5 which show little dependency of these two factors.

*H2: Internal factors such as those with higher rating, closer to the airport, hotel capacity and having a swimming pool are strongly responsible for room rents. Other factors such as free Wifi and free breakfast also contribute a little to the price range of the room rents.*

* 1. **Data**

India is a very well known country for all the tourism that it can attract, for example the hill stations, the Taj Mahal in Agra, the Brindavan gardens in Bangalore etc.. and many more have much of historical relevance as well as aesthetic beauty which attracts tourisits from all over the world. Being as rich as always, India has the ability to attract the tourists with its huge range of hotels right from 1 star to 7 stars and much more with different range of prices. Our analysis here is about the variation in prices and what are the factors that drive to bring in those variations in the price ranges. Our data here encompasses 42 cities which includes – Agra, Ahmedabad, Amritsar, Bangalore, Bhubaneswar, Chandigarh, Chennai ,Darjeeling, Delhi,

Gangok, Goa Guwahati ,Haridwar ,Hyderabad ,Indore ,Jaipur ,Jaisalmer ,Jodhpur ,Kanpur,

Kochi, Kolkata ,Lucknow ,Madurai,Manali ,Mangalore ,Mumbai ,Munnar,Mysore ,Nainital ,Ooty ,Panchkula ,Pune,Puri ,Rajkot,Rishikesh,Shimla ,Srinagar,Surat,Thiruvanthipuram ,

Thrissur,Udaipur and Varanasi.

Now let us consider two cities to understand what exactly the data set is trying to tell us. Let us consider Mumbai and Bhubaneshwar –

Mumbai – This is a metropolitan city that is ranked first and is a tourist destination as well. The stats regarding the weekend and whether it is new year’s eve have a p value of less than 0.5 but have very little significance only so we can move on to the next few rows. We do observe that room rents are highest for rooms with a star rating of 5 with a maximum value of Rs. 24378. The distance from the airport is also pretty decent and the hotel capacity is really large in hotels that have a higher star rating. Then we do observe a few categorical columns that do help us to find out about the availabitly of free wifi, free breakfast and a swimming pool. We do understand from the summary statistics of Mumbai that free wifi and free breakfast are mostly provided in most of the hotels. Although, a mean of 0.345 is the availability of a swimming pool.

Contrast to the above example, let us take a look at Bhubaneshwar. It has a ranking of 28th place. And is neither a metropolitan city nor a tourist spot. The average price is Rs. 2462 and is relatively cheaper than other room rents. All the hotels have free Wi-Fi and few have free breakfast and mostly none have swimming pools. The hotel capacity averages up to 27.87. Therefore, data could be analysed by understanding the columns accordingly.

It is indeed probable that many factors govern the prices of hotel rooms all around India. Any meaningful empirical analysis will need to control for factors. For example, factors such as the city the hotel is located in, whether it is a tourist destination, the hotel rating, its capacity, how far is it from the airport and other features like free Wi-Fi, breakfast and having a swimming pool are all likely to influence hotel prices.

**City Rank :** This is a dummy variable that we use to denote the rankings related to various cities indicated in the list. This data is collected from the official website as per 2016 norms and the data set starts counting from 0 to 44 to show the ranks. It upholds 0 value for Mumbai and 44th value for Manali.

**IsMetroCity:** This is a dummy variable used with categorical reference. The value 0 states that it is not a metropolitan city and 1 indicates that it is one. The 1 value is given to all of the four mentioned cities i.e Mumbai,Chennai, Delhi and Kolkata. All other cities will have a value of 0 for this column. Metropolitan cities are referred to those that have a well defined railway and airport services with developments ahead of other cities.

**IsTouristDestination:** Not every place is a tourist spot. Those locations that possess monuments, museums and other fascinating spots come under this category. Therefore, it is a common conception that people visit tourist spots more frequently than any other place. Now, in places where the tourists population is high, the room rents also go higher. Again, here we use a dummy variable to denote 1 for a tourist destination and 0 for a city that isn’t designed for tourism.

**Star Rating:** This is a simple number representation in the data set. According to the Ministry a scheme for classification of operational hotels using a Star rating has been formulated many years ago. 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 strongly positive correlation with the price of the hotel rooms. Hence, price variations have to be analysed by taking the star ratings into consideration.

**Airport:** This is again a number representation that depicts the distance between the hotel and the nearest airport in kilometres.This is a very important factor to take into consideration. It helps us to analyse the distance between the hotel and the airport. For ease of access and many other reasons people prefer to stay closer to the airport. Also, for officials who lack enough time, they prefer hotels near the airport and an increase in prices can be observed if the hotel room in closer to the airport.

**Free Wi-Fi and Free Breakfast:** These are two dummy variables which are categorical in nature and they tell us whether the hotel rooms have free Wi-Fi with the 1 variable or no free Wi-Fi with a value of 0.Similarly, 1 for free breakfast and 0 for no free breakfast. These have light significance to the room rent as places which have free Wi-Fi and breakfast charge a huger amount relatively but with very small variation only.

**Has Swimming Pool:** It is a dummy variable once again which denotes 1 for having a swimming pool and 0 for not having one. This is also a deciding factor as usually hotels with a good star rating only possess swimming pools and hence an increase in the price. Therefore the variation of this variable is important for room rent analysis.

**Hotel Capacity:** This denotes the amount of rooms that each hotel possess which can in turn tell us about the occupancy rate in the hotels. This is a number representation which tells us the exact number of rooms the hotel has. This factor also has a huge correlation with the room rent factor.

**Room Rent:** This is the price value in rupees that tells us the exact amount of money involved in renting a particular room with respect to the specifications of the other columns.

**Summary Statistics of the hotel pricing in India (Mean Values)**

**CityName RoomRent IsTouristDestination StarRating Airport HotelCapacity**

1 **Agra** 4124.287 1.0000000 3.416667 7.927778 56.61

2 **Ahmedabad** 4175.045 0.0000000 3.481132 8.479245 71.47

3 **Amritsar** 3444.029 1.0000000 3.352941 3.876471 52.11

4 **Bangalore** 4112.803 0.0000000 3.707317 29.923171 86.92

5 **Bhubaneswar** 3587.442 0.0000000 3.200000 4.566667 27.86

6 **Chandigarh** 4030.940 0.0000000 3.488095 6.242857 55.11

7 **Chennai** 4323.647 0.2115385 3.730769 11.261538 117.53

8 **Darjeeling** 5458.088 1.0000000 3.264706 35.329412 31.29

9 **Delhi** 4318.606 1.0000000 3.572656 14.696484 82.69

10 **Gangtok** 4629.648 1.0000000 3.250000 124.000000 27.93

11 **Goa** 8170.801 1.0000000 3.294872 20.928205 48.88

12 **Guwahati** 5325.812 1.0000000 3.500000 25.233333 52.66

13 **Haridwar** 3919.938 1.0000000 3.583333 20.000000 56.66

14 **Hyderabad** 3852.175 0.0000000 3.477612 20.921642 80.39

15 **Indore** 3414.594 0.0000000 3.200000 11.340000 43.55

16 **Jaipur** 7292.022 1.0000000 3.541667 10.495313 61.46

17 **Jaisalmer** 5986.072 1.0000000 3.075758 6.609091 29.48

18 **Jodhpur** 10661.371 1.0000000 3.410714 6.646429 37.57

19 **Kanpur** 3008.562 0.0000000 3.000000 9.700000 10.00

20 **Kochi** 6039.609 1.0000000 3.421053 24.109211 47.46

21 **Kolkata** 4528.986 0.3613281 3.404297 12.683398 58.51

22 **Lucknow** 5879.070 0.0000000 3.468750 11.231250 59.75

23 **Madurai** 4768.223 1.0000000 3.214286 14.000000 37.14

24 **Manali** 4858.285 1.0000000 3.333333 39.953472 32.19

25 **Mangalore** 4110.337 0.0000000 3.307692 11.323077 46.53

26 **Mumbai** 6343.730 1.0000000 3.606742 12.643820 102.70

27 **Munnar** 7543.500 1.0000000 3.353659 106.007317 30.50

28 **Mysore** 3320.869 1.0000000 3.125000 14.335000 52.75

29 **Nainital** 6409.833 1.0000000 3.416667 65.000000 30.33

30 **Ooty** 6144.257 1.0000000 3.294118 100.000000 49.23

31 **Panchkula** 2813.500 0.0000000 3.250000 10.800000 45.37

32 **Pune** 3897.652 0.0000000 3.430667 9.889333 79.60

33 **Puri** 5708.429 1.0000000 3.142857 62.428571 42.14

34 **Rajkot** 4107.078 0.0000000 3.437500 3.612500 29.37

35 **Rishikesh** 4943.670 1.0000000 3.227273 28.672727 27.81

36 **Shimla** 5780.604 1.0000000 3.285714 35.540000 40.20

37 **Srinagar** 10572.025 1.0000000 4.200000 27.780000 43.20

38 **Surat** 3660.850 0.0000000 3.400000 15.900000 56.90

39 **Thiruvanthipuram** 6726.796 1.0000000 3.744898 11.718367 49.73

40 **Thrissur** 3387.844 1.0000000 3.500000 60.625000 28.00

41 **Udaipur** 10145.252 1.0000000 3.343860 24.708772 41.98

42 **Varanasi** 8675.042 1.0000000 3.248485 22.242424 27.54

The above data helps us to summarize a few parameters of immense importance that affect the price of various hotels in different cities that is analysed with this data set. With this set it is easier to find a consolidated form of all the mean values of the factors.

* 1. **Model**

The analysis of the factors that affect the room rent pricings are done with the help of two models as described below.

**Model 1:**

This model has been formulated with the help of the following few parameters i.e

IsMetroCity, IsTouristDestination, IsWeekend, IsNewYearEve, Date,StarRating, FreeWifi, FreeBreakfast, HotelCapacity, HasSwimmingPool, Airport and City Rank. This regression was done with all the parameters that had some levels of standard deviation as well as some

numerical factor involved.

*RoomRent = b0 + b1\* IsMetroCity +b2\* IsTouristDestination + b3\* IsWeekend + b4\* IsNewYearEve + b5\* Date + b6\* StarRating + b7\* FreeWifi +b8\* FreeBreakfast + b9\* HotelCapacity + b10\* HasSwimmingPool + b11\*Airport + b12\*CityRank*

For this model the Adjusted R-squared is 0.1879 and a p-value: < 2.2e-16, we use regression here to find out the factors that have high significance. With the factors that have high

significance the p value is less than 0.05 and these factors help to reject the null hypothesis.

Now inorder to check the best predictors in the above model – we use the leaps library and create a plot to denote the factors that prove to be a good predictor for the room rent analysis. Therefore, after doing that, the factors which don’t act as good predictors are removed and give rise to the new model.

**Results of Model 1**

**Residuals:**

Min 1Q Median 3Q Max

-11895 -2327 -714 1051 310252

**Coefficients:**

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

(Intercept) -10727.114 1163.519 -9.220 < 2e-16 \*\*\*

IsMetroCity -1173.450 177.719 -6.603 4.19e-11 \*\*\*

CityRank 24.240 6.740 3.596 0.000324 \*\*\*

IsTouristDestination 1813.750 149.916 12.098 < 2e-16 \*\*\*

IsWeekend 382.419 441.790 0.866 0.386719

IsNewYearEve 1771.422 1419.073 1.248 0.211945

Date21-Dec-16 453.169 1474.941 0.307 0.758662

Date24-Dec-16 144.000 1407.222 0.102 0.918497

Date25-Dec-16 -17.205 1407.222 -0.012 0.990246

Date28-Dec-16 465.669 1474.941 0.316 0.752220

Date31-Dec-16 -1567.899 1998.510 -0.785 0.432741

Date4-Jan-16 1741.588 1609.775 1.082 0.279324

Date4-Jan-17 875.245 2129.967 0.411 0.681138

Date8-Jan-16 1528.041 1547.813 0.987 0.323550

Date8-Jan-17 506.288 2083.915 0.243 0.808048

DateDec 18 2016 720.970 1018.420 0.708 0.479001

DateDec 21 2016 1285.550 1103.372 1.165 0.243995

DateDec 24 2016 1384.425 1018.399 1.359 0.174039

DateDec 25 2016 1306.273 1018.399 1.283 0.199629

DateDec 28 2016 1801.738 1104.195 1.632 0.102762

DateDec 31 2016 294.441 1735.780 0.170 0.865304

DateJan 04 2017 1890.441 1103.921 1.712 0.086832 .

DateJan 08 2017 1194.058 1019.181 1.172 0.241385

DateJan 4 2017 1635.404 1392.761 1.174 0.240329

DateJan 8 2017 1020.527 1296.050 0.787 0.431054

StarRating 3567.454 110.454 32.298 < 2e-16 \*\*\*

FreeWifi 547.907 224.302 2.443 0.014590 \*

FreeBreakfast 170.618 123.437 1.382 0.166924

HotelCapacity -10.645 1.029 -10.341 < 2e-16 \*\*\*

HasSwimmingPool 2239.160 159.405 14.047 < 2e-16 \*\*\*

Airport 7.370 3.039 2.425 0.015303 \*

---

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

Residual standard error: 6600 on 13201 degrees of freedom

Multiple R-squared: 0.1917, Adjusted R-squared: 0.1898

F-statistic: 104.3 on 30 and 13201 DF, p-value: < 2.2e-16

**Model 2:**

For this model we eliminated all the factors that are not good predictors and hence we have all values which provide us with proper significance. Now a clearer model is obtained as all the irrelevant data is removed from the analysis.

*RoomRent = b0 + b1\*IsMetroCity + b2\*IsTouristDestination+b3\*StarRating+ b4\*HotelCapacity+ b5\*HasSwimmingPool + b6\*CityRank*

The R value here is smaller than for model 1 which indicates that the model is more precise than the first model. Also the AIC value is lesser for the second model than the first. Therefore by all means, model two fits better than Model 1 for this particular analysis of the room

rent. Akaike Information Criterion (AIC) developed by Akaike, (1974) and the Bayesian Information Criterion (BIC) developed by Schwarz (1978), represent the trade-off between the goodness of fit of the model and the complexity of the model. If Model 2 indeed fits the data better than Model 1, we expected the AIC and BIC of Model 2 to be less than Model 1.

**Results for Model 2:**

**Residuals:**

Min 1Q Median 3Q Max

-11657 -2347 -743 1032 309603

**Coefficients:**

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

(Intercept) -8545.177 356.604 -23.963 < 2e-16 \*\*\*

IsMetroCity -1027.288 172.421 -5.958 2.62e-09 \*\*\*

CityRank 33.307 5.914 5.632 1.82e-08 \*\*\*

IsTouristDestination 1832.210 145.708 12.575 < 2e-16 \*\*\*

StarRating 3595.391 110.122 32.649 < 2e-16 \*\*\*

HotelCapacity -10.781 1.024 -10.527 < 2e-16 \*\*\*

HasSwimmingPool 2184.793 157.762 13.849 < 2e-16 \*\*\*

---

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

Residual standard error: 6611 on 13225 degrees of freedom

Multiple R-squared: 0.1877, Adjusted R-squared: 0.1873

F-statistic: 509.2 on 6 and 13225 DF, p-value: < 2.2e-16

* 1. **Results**

**Model 1:** Hence model 1 has helped us to conclude that our hypothesis are right to a major extent and helps us to prove our hypotheses empirically. This is done by analysing the p value of most of the regressions made. The p-value below 0.01 helps us to say that that particular factor has high significance and that the factor helps to reject the null hypothesis. The coefficients are also greater than 0 which help to say that that factor has high significance.

**Model 2:** The second model was done as an improvisation to the first model. After using the leaps command we found the factors that best fit the model. And in this process the factors that were in model 1 like date, free Wi-Fi, free breakfast, is weekend and is new year’s eve were all removed due to very little correlation. Now this model had an Adjusted r squared value less and AIC value less than Model 1 and hence this model 2 is said to be more fit than Model 1.

This model also helps us to say that our two hypotheses done above are correct and can be depicted by a p-value of less than 0.01 and coefficients which are positive. Therefore, we can say that Metropolitan cities, destination spots and high ranked cities have higher prices of room rents and also factors like having a swimming pool, star rating, airport distance and hotel capacity have high relevance to the rent.

1. **Conclusion**

This paper has provided an extended knowledge of the extent to which a few parameters affect the room rent pricings. Such effects have to be analysed and their extent has to be taken into account. This paper mainly focused on all the factors that are highly correlated with the room rent and helps us to denote the factors that affect the room rent the most. Such a deep analysis is required to make sure that all the hotels set up are actually getting the amount of money that they deserve or are they being over paid.

This paper has high managerial relevance as it allows one to have an in depth study of these factors. Now, this provides managers with insights that help in setting budgets and expecting a profit in return for all the money put into a particular hotel industry. This way the quality of the hotels, value for money and other extra internal factors may be taken into consideration as per the manager’s call. This way a happier supplier – customer relationship can be formed and the hotel industry may see a boom in its carrier.