# Project Title: Report for Lodging price for Tourists

# NAME: Harish Ravi

# EMAIL: [hharish001@gmail.com](mailto:hharish001@gmail.com)

# COLLEGE / COMPANY: SSN college of Engineering

**Lodging price for Tourists, India**

1. **Introduction:**

This paper addresses the room rents for Boarding & Lodging in some of India’s popular hotels located at tourist and the factors that affect it. It uses the dataset provided by [www.hotels.in](http://www.hotels.in) for 42 cities in India. The idea of this paper is to find out the major factors that decide the room rent for hotels in India located in tourist destinations and to find the pricing patterns involved in it

1. **Overview of the study:**

Our study concerns about the hotel pricing in India with regard to tourism and analyze how it varies with each and every factor like city, star rating, capacity , So this is a study which is done on a data collected from hotels in India , So we try to derive some

Useful conclusions which will help us predict the trend of pricing of Indian hotels located at cities that get a fairly good amount of tourists

We do some tests to test out hypothesis etc., We point out some variables which are highly correlated to the pricing of hotel rooms.

We found that hotel pricing is the variable of interest and also we found out by testing the correlations that factors like Star Rating , hotel having a swimming pool ,Hotel Capacity to be some important correlations and some loosely correlated variables are Population of the city visited and if it’s a metro city or not and also the city visited

**3. An empirical study of hotel pricing**

The objective of this study is to find out the correlations between the various factor that affect the pricing of hotel rooms in India and to build a regression model that predicts the room price given the input factors with a statistical degree of accuracy

**3.1 Hypothesis**

With our initial tests for correlation we were able to find some correlations and based on which we put out some correlations

**H1: Hotels with higher star rating charge much higher**

**H2: Hotels that have a swimming pool charge much**

**H3: Hotels that have a higher capacity charge more**

**A corrgram that gives some information on the correlations in data**

****

These hypothesis states that hotels that offer more luxuries (star rating or features like swimming pool) or large hotels (capacity) or is located in a tourist destination is more likely to charge higher than those which do not satisfy these features

**3.2 Data:**

India has a lot of tourist destinations among cities in fact most of our dataset in consideration is a tourist destination with around 9215 hotels in it and 3997 hotels in a non-tourist destination.

And also the data followed the same pattern of correlations as it did before so it’s safe to assume that the majority of these tourists were Indians

**Room Rent:**

As the name denotes it’s the room rent the guests are expected to pay for their stay,

This is the main variable to be predicted in model. Let’s call it y

**Star Rating:**

Star Rating denotes the quality of service offered by a hotel, higher the rating and so is the facilities , Thus we see that star rating is a major factor in pricing. We used this independent variable as a direct measure of the amenities provided by the hotel and hence is a factor of pricing, Let’s call this x1

**Has Swimming pool:**

This variable states whether a hotel has a swimming pool or not , It’s a discrete variable with either 1 or 0 , As we’ve found out that it has a high correlation with the Room rent this is taken in consideration in our model , Let’s call it x2

**Hotel Capacity:**

This variable is used to denote the capacity of the hotel and usually denotes the size and scale of the hotel, As large hotels are known to provide more luxurious services this is a major factor in pricing, let’s call it x3

**Is Metro city:**

This variable specifies if the city under consideration is a metro city or not, It

seems to have some correlation with pricing so we take it into consideration, let’s call it x4

**Population of the city:**

This variable denotes the population of the city under consideration, as large cities are metropolitan and have more amenities this is taken into consideration, let’s call it x5

**City:**

This is the city in which the survey was conducted, as evidenced from the data the city does have an effect on the room rent

|  |  |  |
| --- | --- | --- |
|  |  | |
| S.no | City name | **Mean Room rent** |
| 1 | Agra | 4124.287 |
| 2 | Amritsar | 3444.029 |
| 3 | Chennai | 4009.341 |
| 4 | Darjeeling | 5458.088 |
| 5 | Delhi | 4318.606 |
| 6 | Gangtok | 4629.648 |
| 7 | Goa | 8170.801 |
| 8 | Guwahati | 5325.812 |
| 9 | Haridwar | 3919.938 |
| 10 | Jaipur | 7292.022 |
| 11 | Jaisalmer | 5986.072 |
| 12 | Jodhpur | 10661.371 |
| 13 | Kochi | 6039.609 |
| 14 | Kolkata | 3899.659 |
| 15 | Madurai | 4768.223 |
| 16 | Manali | 4858.285 |
| 17 | Mumbai | 6343.730 |
| 18 | Munnar | 7543.500 |
| 19 | Mysore | 3320.869 |
| 20 | Nainital | 6409.833 |
| 21 | Ooty | 6144.257 |
| 22 | Puri | 5708.429 |
| 23 | Rishikesh | 4943.670 |
| 24 | Shimla | 5780.604 |
| 25 | Srinagar | 10572.025 |
| 26 | Thiruvanthipuram | 6726.796 |
| 27 | Thrissur | 3387.844 |
| 28 | Udaipur | 10145.252 |
| 29 | Varanasi | 8675.042 |

The correlations for the hypothesis are as follows

|  |  |
| --- | --- |
| Hypothesis | Correlation |
| H1 | 0.3782 |
| H2 | 0.3106 |
| H3 | 0.1463 |

Thus we identify that the Star rating , swimming pool and hotel capacity are three most important factors that determine pricing

**3.3 MODEL:**

**Model 1:**

We first established that Star Rating, Room Rent decide the pricing of rooms more than any other factors

This model is fairly straightforward in predicting the room rent

Where

y = (1)

where x1 is the Star Rating

x2 – has swimming pool

and y is the Room Rent

Model 1:

|  |  |  |
| --- | --- | --- |
| Factor | **β** | SE |
| Intercept | -6123.5\*\*\* | 414.7 |
| Star Rating | 3288.4\*\*\* | 207.5 |
| Has swimming pool | 2388.2\*\*\* | 207.5 |

AIC = 191526.7

**Model 2 :**

Now consider Star Rating , Room Rent and hotel capacity deciding the pricing of rooms

y = (2)

here x3 is the Hotel Capacity

Model 2:

|  |  |  |
| --- | --- | --- |
| Factor | **β** | SE |
| Intercept | -8042.025\*\*\* | 440.033 |
| Star Rating | 4067.060\*\*\* | 143.618 |
| Has Swimming pool | 2857.806\*\*\* | 209.340 |
| Hotel Capacity | -16.145 | 1.315 |

AIC = 191379

The AIC value for this model is higher than model 1 on addition of a variable

**Model 3:**

It further improved to 191147.4

When variables like City, Population of the city, Is metro city were further added

y = (3)

here x4 is the city name

x5 is the population

x6 is if it’s a metro city or not

Model 3:

|  |  |  |
| --- | --- | --- |
| Factor | **β** | SE |
| Intercept | -8952 | 463.2 \*\*\* |
| Has swimming pool | 2640 | 210 \*\*\* |
| Star Rating | 3986 | 142.3 \*\*\* |
| Hotel Capacity | -11.78 | 1.347 \*\*\* |
| City name | 75.38 | 7.751 \*\*\* |
| Population | 4.672 x 10ˆ-5 | 4.382 x 10 ^ -5 |
| Is metro city | -1933 | 426.2 \*\*\* |

Model 3 seemed to have better accuracy over model 1 or model 2

**4.Conclusion:**

The paper was motivated to study the factors involving hotel pricing in India, From model 3 it was clear that even though population had correlation with room rent it had no role in regression ,

Similarly the addition of any more variables for predictions did not improve the result obtained, This does have some managerial implications as well such as deciding when and where hotels started could yield more profit and factors for it to pay off more , and also for example deciding the perfect city to start a good hotel and what is the good room rent for the given features for that hotel in that city .

This model helps identify the features for which people pay more , what they don’t care about and what they expect

**Figure 1: Rambagh palace in Jaipur a 5-star hotel**

