1. **Introduction**

The room rent for various hotels in various cities on eight different days are collected. All the factors which help recognize a hotel distinctly and which affect or influence the room rent in some way are considered as variables in this particular dataframe and the corresponding values are noted down.

This paper addresses the factors which affect the Room rent of a particular hotel.

Does a swimming pool in a hotel affect the rent?

Do hotels in Metro cities have a higher rent?

If the hotel is located in a city which is a tourist destination, will it affect the rent?

We find the answers to all such questions and much more through this analysis.

2. **Overview of the Study**

Our field of study is all the hotels located in Indian cities.

The data was collected from [www.hotels.in](http://www.hotels.in) in October 2016.

The dataset tracks hotel prices on 8 different dates at different hotels across different cities.

3. **Data:**

Certain hotels are picked up from Indian cities and their pricing strategy is studied in detail. We have collected several types of data for each hotel so that we can recognize each hotel distinctly and also analyse the factors that are contributing to the Room rent of all these hotels.

**Notice that the dataset tracks hotel prices on 8 different dates at different hotels across different cities.**

**The data was collected from** [**www.hotels.in**](http://www.hotels.in) **in October 2016.**

There are several factors on which data is collected and noted down.

1. RoomRent

Units : Rupees

**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.**

1. Date

Units : text

**We have hotel room rent data for the following 8 dates for each hotel:**

**{Dec 31, Dec 25, Dec 24, Dec 18, Dec 21, Dec 28, Jan 4, Jan 8}**

**If a hotel is sold out on a given date, assume that the price of the hotel room on the date it is sold out is the maximum price from the sample of dates for which prices are available.**

1. IsWeekend

Units: Dummy variable

**We use ‘0’ to indicate week days, ‘1’ to indicate weekend dates (Sat / Sun)**

1. IsTouristDestination

Units: Dummy variable

**We use ‘1’ if the city is primarily a tourist destination, ‘0’ otherwise. For example, Goa and Agra are primarily tourist destinations. We assume that most people who visit Goa and Agra and stay in their hotels are in these cities primarily for tourism.**

1. IsNewYearEve

Units: Dummy variable

**1’ for Dec 31, ‘0’ otherwise**

1. HasSwimmingPool

Units: Dummy variable

**‘1’ if they have a swimming pool, ‘0’ otherwise**

1. FreeBreakfast

Units: Dummy variable

**‘1’ if the hotel offers Free Breakfast, ‘0’ otherwise**

1. FreeWifi

Units: Dummy variable

**‘1’ if the hotel offers Free Wifi, ‘0’ otherwise**

1. IsMetroCity

Units: Dummy variable

**‘1’ if CityName is {Mumbai, Delhi, Kolkatta, Chennai}, ‘0’ otherwise**

1. CityRank

Units: Dummy variable

**Rank order of City by Population (e.g. Mumbai = 0, Delhi = 1, so on)**

1. HotelCapacity

Units: Number

**Number of people who can be accommodated. e.g. 242.  (enter ‘0’ if not available)**

1. HotelPincode

Units: Number

**Pincode of the hotel eg :403712**

1. StarRating

Units: Number

**Number of stars for the hotel( Number can vary from 5 to 1)**

1. HotelAddress

Units: text

**Precise location of that hotel along with city name**

1. HotelDescription

Units: text

**e.g. 5-star beachfront resort with spa, near Arossim Beach**

1. HotelName

Units: text

**The name of the hotel e.g. Park Hyatt Goa Resort and Spa**

1. Airport

Units: kilometers

**Distance from the hotel to the nearest airport**

1. Population

Units: Number

**Population of the City in 2011**

1. CityName

Units: text

**Name of the city**

**Table A1:  City Rank (based on 2011 City Population)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CITYRANK** | **CITYNAME** | **IsHolidayDestination** | **Number of Hotels listed on Hotels.com** | [**City Population(2011)**](https://en.wikipedia.org/wiki/List_of_cities_in_India_by_population#cite_note-Cities1Lakhandabove-3) |
| 0 | Mumbai | 0 | 405 | 12,442,373 |
| 1 | Delhi | 0 | 871 | 11,034,555 |
| 2 | Bangalore | 0 | 450 | 8,443,675 |
| 3 | Chennai | 0 | 287 | 7,088,000 |
| 4 | Hyderabad | 0 | 237 | 6,731,790 |
| 5 | Ahmedabad | 0 | 136 | 5,577,940 |
| 6 | Kolkata | 0 | 192 | 4,496,694 |
| 7 | Surat | 0 | 20 | 4,467,797 |
| 8 | Pune | 0 | 205 | 3,124,458 |
| 9 | Jaipur | 1 | 286 | 3,046,163 |
| 10 | Thrissur | 0 | 36 | 2,975,440 |
| 11 | Lucknow | 0 | 37 | 2,817,105 |
| 12 | Kanpur | 0 | 13 | 2,765,348 |
| 13 | Amritsar | 1 | 72 | 2,490,891 |
| 14 | Indore | 0 | 49 | 1,960,631 |
| 15 | Agra | 1 | 102 | 1,760,285 |
| 16 | Madurai | 1 | 21 | 1,465,625 |
| 17 | Goa | 1 | 626 | 1,457,723 |
| 18 | Rajkot | 0 | 26 | 1,286,678 |
| 19 | Varanasi | 1 | 60 | 1,201,815 |
| 20 | Srinagar | 1 | 57 | 1,180,570 |
| 21 | Jodhpur | 1 | 81 | 1,033,918 |
| 22 | Chandigarh | 0 | 117 | 960,787 |
| 23 | Thiruvathipuram | 0 | 128 | 957,730 |
| 24 | Guwahati | 0 | 12 | 957,352 |
| 25 | Mysore | 1 | 58 | 887,446 |
| 26 | Bhubaneswar | 0 | 29 | 837,737 |
| 27 | Kochi | 1 | 188 | 595,575 |
| 28 | Mangalore | 0 | 13 | 499,487 |
| 29 | Udaipur | 1 | 113 | 451,735 |
| 30 | Pondicherry | 0 | 42 | 241,773 |
| 31 | Haridwar | 1 | 73 | 228,832 |
| 32 | Puri | 1 | 24 | 201,026 |
| 33 | Shimla | 1 | 58 | 169,578 |
| 34 | Panchkula | 0 | 118 | 140,925 |
| 35 | Darjeeling | 1 | 32 | 132,016 |
| 36 | Rishikesh | 1 | 107 | 102,138 |
| 37 | Gangtok | 1 | 30 | 98,658 |
| 38 | Ooty | 1 | 64 | 88,430 |
| 39 | Jaisalmer | 1 | 82 | 65,471 |
| 40 | Nainital | 1 | 85 | 41,377 |
| 41 | Munnar | 1 | 108 | 38,471 |
| 42 | Manali | 1 | 80 | 8,096 |

4. **T tests for categorical variables**

**Please refer to the Output document for detailed analysis. Just the results are summarized below.**

**Hypothesis 1**

1. IsWeekend

H0: The room rent of hotels on weekends and weekdays are the same

H1: The room rent of hotels on weekends are more than on weekdays.

The p value is more than 0.05 so it does not have a significant impact on the room rent. We fail to reject null hypothesis.

**Hypothesis 2**

1. FreeWifi

H0: The room rent of hotels having free wifi and room rent of hotels without free wifi are almost similar

H1: The room rent of hotels having free wifi is greater than the room rent of hotels without free wifi

The p value is more than 0.05 so it does not have a significant impact on the room rent. We fail to reject the null hypothesis.

**Hypothesis 3**

1. FreeBreakfast

H0: The room rent for hotels having free breakfast and for hotels without free breakfast is the same.

H1: The room rent for hotels having free breakfast is more than the room rent for hotels without free breakfast

The p value is more than 0.05 so it does not have a significant impact on the room rent. We fail to reject null hypothesis.

**Hypothesis 4**

d. IsNewYearEve

H0: The room rent for hotels on New year eve and on other days is the same

H1: The room rent for hotels on New year eve is more than the room rent on the other days.

The p value is less than 0.05 so the room rent of hotels on new year eve is more than the other days.

We reject null hypothesis.

**Hypothesis 5**

1. IsMetroCity

H0: The room rent of hotels in metro cities and non metro cities is the same

H1: The room rent of hotels in metro cities is more than the room rent of hotels in non metro cities

The p value is less than 0.05 so the room rent of hotels in metro cities is more as compared to non metro cities.

We reject the null hypothesis.

**Hypothesis 6**

f. IsTouristDestination

H0: The room rent of hotels in tourist cities and non tourist cities is the same

H1: The room rent of hotels in tourist cities is more than room rent of hotels in non tourist cities

The p value is less than 0.05 so the room rent of hotels in cities which are tourist destinations is more than the room rent of hotels in cities which are not tourist destinations.

We reject the null hypothesis.

**Hypothesis 7**

g. HasSwimmingPool

H0: The room rent of hotels having swimming pool and without swimming pool is the same

H1: The room rent of hotels having swimming pool is more than the room rent of hotels without swimming pool

The p value is less than 0.05 so the room rent of hotels having swimming pool is more than the hotels without swimming pool.

We reject the null hypothesis.

***So we can conclude from the t tests that IsMetroCity, IsTouristDestination, IsNewYearEve and HasSwimmingPool have p values less than 0.05 and are statistically significant. Hence they play an important role in determining the Room rent of a particular hotel.***

5. **LINEAR REGRESSION MODELS**

Model 1:

For the first model we consider most of the variables in the data set.

We ignore the cityrank because it is already covered in the city name, and date because it is covered in IsNewYearEve

We also exclude the Hotel name, address,pincode and description as they are specific to a particular hotel and cannot be interpreted as factors which affect the room rent.

We notice that the variables such as CityName, Population, IsMetroCity, IsTouristDestination,

IsWeekend, FreeBreakfast are not statistically significant as indicated by the lack of stars on

the right hand side.

So while constructing the next model, we eliminate the above mentioned variables.

Model 2:

For the model 2, we eliminate the variables CityName, Population, IsMetroCity, IsTouristDestination, IsWeekend, FreeBreakfast.

We notice that FreeWifi is not statistically significant as indicated by the lack of

Stars. So while building the next model, we eleiminate the FreeWifi variable.

Model 3:

For the model 3, we eliminate the FreeWifi variable too. Now all the variables show a high level of significance( three stars).

We find that model 3 is the best fit model, since all the independent variables are statistically significant.

***Based on the regression analysis we conclude that the IsNewYearEve, StarRating, Airport, HotelCapacity, HasSwimmingPool are the most important independent variables to determine the RoomRent of a particular hotel.***

**Regression model equation**

**RoomRent = (835.597)IsNewYearEve + (3521.996)StarRating + (25.337)Airport + (-14.775) HotelCapacity + (2708.650) HasSwimmingPool + (-7388.523)**

6. **Conclusion**

This paper was motivated by the need for research that could improve our understanding of how various factors contribute to the Room rent of hotels in different cities.

**Based on the regression analysis we conclude that the IsNewYearEve, StarRating, Airport, HotelCapacity, HasSwimmingPool are the most important independent variables to determine the RoomRent of a particular hotel.**