**Hotel Room Pricing In The Indian Market**

1. **Introduction**

Over the past decade or so, India has seen a surge in tourism. Our vast and varied culture attracts tourists from various parts of the world. Foreigners visit our country to get a taste of our heritage while locals go on vacations with family and friends. This surge in tourism has been profitable to 1 industry- the hotel industry. According to an article in Business Insider, the hotel industry already accounts for 7.5 per cent of the country's GDP. According to a report by KPMG, the hospitality sector in India is expected to grow at 16.1 per cent CAGR to reach Rs 2,796.9 thousand crore in 2022 (<http://businessworld.in/article/Hospitality-Industry-In-India-A-Big-Contributor-To-Economy-s-Growth-/16-05-2017-118291/>).

Majority of revenue earned by hotels is through renting rooms. In our project we will analyse the factors that determine the price of a hotel room per day.

**2. Overview of the study.**

Our field of study concerns hotel room pricing in India and the factors that affect these prices. We analyse what factors affect the price of hotel rooms and what affect they have. Our regression analysis reveals that the price can be determined with the following factors:

* Date of stay
* Star rating of hotel
* Is the city a tourist destination/metro
* Is it a weekend
* Is it New Year’s eve
* Name of the hotel and city
* Features(Free wifi and breakfast, swimming pool)

**3. An empirical field study of hotel room prices in India**

**3.1 Hypothesis**

We study how the price of the room is correlated with various facilities and location of the hotel and even the brand (name ) of the hotel .We expect that the rooms with facilities like Free Wifi,Free Breakfast and near to airport, near tourist destination, city rank and on various holiday and some hotels brand(names ) like Taj,Marriott,ITC have higher room rent.

Therefore we, make following hypothesis

**H1: The room rent are generally high for the rooms with the following facilities and based on location following location factors and the star ratings**

**3.2 Data**

For this project, the data was collected from [www.hotels.in](http://www.hotels.in). The dataset had the following fields

**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.

**Date**

The hotel business runs based on a “seasonal” model. There is a time of the year where tourism is high and prices surge. There are times when holidays are scarce, climate is unsuitable for tourism. This is when prices decline. 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}

**IsWeekend**

Day of the week can affect prices. If a long weekend is in the horizon, more people will plan holidays and outings which will lead to surge in the pricing. Weekdays tend to not attract high prices. In our database we use a dummy variable which holds a binary value

0-weekday

1-weekend

**IsNewYearEve**

New Year is one of the few festivals that is celebrated by people across the globe. As such, it is more likely to attract tourists and high prices especially if it is New Year’s eve. We use a dummy variable which holds a binary value

0- Not new year’s eve

1- Is new year’s eve

**CityName**

The location of the hotel is a key factor in pricing. If you’re located in one of the big cities like Chennai, Mumbai or Delhi, you are more likely to attract tourists than other cities. These cities are huge and have various tourist attractions which attract people from different places.

**IsMetroCity**

Metro cities are huge. They have everything a big city needs to have. They have sporting activities, hotels,tourist attractions, hospitals, schools, theatres and basically everything which attracts tourists and generates income. We use a dummy variable which holds a binary value

0- City is not metro

1- City is metro

**IsTouristDestination**

It is a dummy variable which contains a binary value. 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.

**HotelName**

If you’re staying at the Taj, you’re going to have to shell out more cash. This is due to branding and brand value. These are hotels which are prefered by guests as they provide various services for the comfort of the customer.

**StarRating**

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 used the variable to denote the star rating of hotel in city

**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 used the variables and to denote the distance of hotel in city from the closest airport and train station respectively.

**HotelAddress**

It gives us the address of the hotel. e.g. Arrossim Beach, Cansaulim, Goa

**HotelPincode**

It is a numeric field. It gives us the hotel pincode. Eg.403712

**HotelDescription**

It gives us a brief description about the hotel and its features. e.g. 5-star beachfront resort with spa, near Arossim Beach

**FreeWifi**

Free wifi is one of the features that are offered to customers to persuade them to stay at hotels.

It is a dummy variable which holds a binary value

0- No free wifi

1- Free wifi is available

**FreeBreakfast**

Since tourists are unlikely to bring cooking utensils, hotels offer free or complimentary breakfast to their guests. Tourists, hence choose hotels which offer this service due to convenience. We use a dummy variable which holds a binary value

0- No free breakfast

1- Free breakfast is provided

**HotelCapacity**

Hotel capacity represents the total number of guests a hotel can accommodate in its rooms. If the number of reservations are less, prices drop to attract more customers. If the number of accommodations left are less, prices are increased to generate maximum revenue.

It is a numeric value which tells us the total capacity of the hotel.e.g. 242. (enter ‘0’ if not available)

**HasSwimmingPool**

It is a dummy variable which holds a binary value

0- No swimming pool

1- Hotel has swimming pool

**3.3 Model**

To determine the prices of the room, we developed a model that considered the following factors

* Date of stay
* Star rating of hotel
* Is the city a tourist destination/metro
* Is it a weekend
* Is it New Year’s eve
* Name of the hotel and city
* Features(Free wifi and breakfast, swimming pool)

These factors were considered as it had the maximum effect on room pricing.

**3.4 Results**

To test our hypothesis, we developed a linear regression model that determined the price of the room based on the factors mentioned above. The model achieved a multiple R-squared value of 0.89 and an adjusted R-squared value of 0.8738 (0- worst fit, 1- perfect fit).

**4. Conclusion**

This project was selected to understand the various factors that affect hotel room pricing and understand the effects each factor had on the final price. By analysing these factors, we were able to build a model that could predict the prices with an accuracy of 89%. This model can be used by tourism agencies, hotel management agencies and others who would be interested in knowing the prices of hotel rooms at any given time.