Saturday, March 18, 2023 2:41 PM

Analysis on resort vs city hotel

**Data Visualization (MGSC-5127-13)**

## SUBMITTED TO:

**Professor Enayat Rajabi**

## SUBMITTED BY:

**Syeda Mubashira 0261612**

**Dharma Devi Udayakumar 0274871 Mariya Dubbaka 0260434**

**Amritpal Kaur 0272882**

# Project Description:

A hotel has a decent chance of becoming prosperous and successful if it is situated in a desirable area and provides top-notch services. Making a profit is one of the main reasons for starting a hotel business. The analysis for resort with city hotel is performed in the aspects of revenue between the two hotels. The factors which are affecting the revenue such as distribution channel, market segment and parking lot. Additionally, the average wait time between two hotels and the guest visits from various European nations are explored.

# Data Set Description:

The data contains total 32 fileds and 11930 records.

We combined a second dataset with a list of countries' full names and their abbreviations.

# Meta Data:

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Variable** | **Data type** | **Description** |
| 1 | hotel | Character | Name of the hotel (Resort Hotel or City Hotel) |
| 2 | is canceled | Integer | If the booking was cancelled (1) or not(0) |
| 3 | Lead time | Integer | Number of days that elapsed between the entering date of the booking into the PMS( property management system) and the arrival date |
| 4 | Arrival date year | Integer | Year of arrival date |
| 5 | Arrival date month | Character | Month of arrival date |
| 6 | Arrival date week number | Integer | week number of year for arrival date |
| 7 | Arrival date day of month | Integer | day of arrival date |
| 8 | Stays in weekend nights | Integer | Number of weekends nights ( Saturday and Sunday)the guest stayed or booked to stay at the hotel |
| 9 | Stays in week nights | Integer | Number of weeknights(Monday to Friday) spent at the hotel |
| 10 | adults | Integer | Number of adults |
| 11 | children | Integer | Number of children |
| 12 | babies | Integer | Number of babies |
| 13 | meal | Character | Type of meal booked. Categories are presented in standard hospitality meal packages : |
|  |  |  | BB – bed and breakfast |
|  |  |  | HB – half board (breakfast and one other meal- usually dinner) |
|  |  |  | FB – full board (breakfast, lunch and dinner) |
| 14 | country | Character | Country of origin |
| 15 | Market segment | Character | Market segment designation. In categories, the term TA means travel agents and TO means tour operators |
| 16 | Distribution channel | Character | Booking distribution channel |
| 17 | Is repeated guest | Integer | if the booking was from a repeated guest (1) or not (0) |
| 18 | Previous cancellations | Integer | Number of previous bookings that were cancelled by the customer to the current booking |
| 19 | Previous booking not cancelled | Integer | Number of previous bookings not cancelled by the customer prior to the current booking |
| 20 | Reserved room type | Character | Code of room type reserved |
| 21 | Assigned room type | Character | Code of room type assigned. Sometimes the assigned room type differs from the reserved room type due to hotel operation reasons or by customer request. |
| 22 | Booking changes | Integer | Number of changes made to the booking |
| 23 | Deposit type | Character | Type of the deposit made by the guest. |
|  |  |  | Non deposit- no deposit was made. |
|  |  |  | Non refund- a deposit was made with a value under the total cost of stay. |
| 24 | Agent | Character | ID of travel agent who made the booking |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 25 | Company | | Character | ID of the company that made the booking | |
| 26 | Days in waiting list | | Integer | Number of days the booking was in the waiting list | |
| 27 | Customer type | | Character | Type of customer assuming one of four categories | |
|  |  | |  | Contract- when the booking has an allotment or other type of contract associated to it. | |
|  |  | |  | Group- when the booking is associated with to a group. | |
|  |  | |  | Transient- when the booking is not part of a group or contract ,and is not associated to other transient booking | |
|  |  | |  | Transient – when the booking is transient , but is associated to at least other transient booking. | |
| 28 | Adr | | Integer | Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights | |
| 29 | Required car parking spaces | | Integer | Number of car parking spaces required by the customer | |
| 30 | Total of special requests | | Integer | Number of special requests made by the customer | |
| 31 | Reservation status | | Character | Reservation status (Cancelled, Check out or No Show) | |
|  |  | |  | Canceled- booking was cancelled by the customer | |
|  |  | |  | Check- out- customer has checked In but already departed | |
| Data | set link: | [https://www.kaggle.co](https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand) | [m/datasets/](https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand) | No show – customer did not check in and did inform the  [jheosstelmoof stthipearke/ahsootnel-wbhoyoking-demand](https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand) | |
|  |  |  |  |
| 32 | Reservation status date | | Integer | Date at which the last reservation status was updated | |

**Data cleaning:**

For an accurate visualisation of the questions, the untidy dataset needs to be cleaned and prepared. The following adjustments were made throughout the cleaning process, for which we used the Tableau Prep tool:

* Removed field “company” as it had 93% NULL values.
* Agent filed had 13% NULL values, replaced it with mean value.
* Removed two records where we had special characters (e.g., hotel – Re$0rt hotel and C!TY hotel)
* Removed the NULL values from below fields

Arrival date month, country, meal, arrival date year.

* We have created two calculated fields.
  1. Total\_ duration = stays \_in\_ weekend\_nights+ stays \_in\_ week\_nights
  2. Revenue = Adr\* Total\_ duration (Revenue indicates total revenue as per reserved date)
  3. Guest count = adults+ children+ babies

updated Clean 1 Output

Changes (14) Filter hotel Exclude : "CITY Hotel" Filter hotel Exclude: Filter arrival\_date\_year Exclude: null Filter arrival date month Exclude: null Group Values children null replaced by "0" Filter meal

Exclude: null Filter country Exclude: null Group Values agent null replaced by "87' Filter adr Exclude. •

-6.38

Changes (14) Filter meal Exclude: null Filter country Exclude: null Group Values agent null replaced by "87' Filter adr Exclude. • -6.38 Filter hotel " ReSOrt Hotel" Exclude: Calculated Field Total duration Calculated Field Revenue Cadr)k [Total \_duration] Calculated Field total number of customers Filter company Exclude: null

Combined both datasets (CleanedOutput2.cvs and CleanedOutput.cvs) using the ‘Join’ function

CleanedOutput2.csv CleanedO... — CleanedO... How do relationships differ from joins? Learn more

CleanedOutput.csv CleanedOutput.csv+ country (CleanedOutpL CleanedOutput2.csv Abc Country Add more fields Performance Options > Operator CleanedOutput.csv Abc country (Cleanec 5 Dashboard 1 Dashboard 2. Dashboard

3. Dashboard 4 AFG ALB DZA ASM AND AGO AIA Dashboard 5

# KEY QUESTIONS:

1. Which hotel is generating more revenue? **(Revenue-based Analysis)**
2. What are the elements that influence the hotel revenue? **(Analysis Using the elements as a basis)**
3. What is the average waiting time for city hotel and resort hotel? **(Analysis of Average waiting time)**
4. Which hotel is the top choice among the countries? **( demographical Analysis of preferred visitors)**

# Design:

We have used tableau desktop for the visualization of our questions. Tableau provided different graphs that could be used to visualize each of our objectives and then we selected the most suitable one among them.

* For first question, we chose the pie chart since it is the best chart for comparing various factors. Here, we're contrasting how much revenue Resort and City hotel generates.
* Second question, since it offers a quick visual summary of information, a heat map is used. The user can comprehend complex data sets with the help of more complex heat maps. According to the investigation, customer type, distribution channel, and meal type are the main variables that affect revenue generation.
* For third question, the Line Graph is selected because the average waiting time over the same period is compared between City hotel and Resort hotel.
* Last question makes use of a choropleth map to display a geographical feature, such as the country where the comparison is done and the proportion of visitors from different European nations to Resort and City hotels are displayed.

**Tableau Visualization:**

# Question1:

The color in the pie chart above denotes the type of hotel. The revenue generated from each hotel category is represented in percentages. 40% of the revenue made by the Resort hotel, and 60% of the revenue made by the City hotel.

# Question 2:

What are the elements that influence the hotel revenue? Distribution Channel hotel ty (All) City Hotel Resort Hotel Custom.. Contract Group Meal sc sc Corporate 2,464 382,862 438 1,003 211,905 3,184 890 Direct GDS 588 96 12,669 357 558 1,748,928 216 143,215

70,336 177,090 14,554 5,578 TA/TO 570,633 5,453 56,147 40,772 2,739 4,453 1,210,292 2,685,741 2, 796,645 9,026 871,413 104,254

Undefined Revenue 14,098,473 Transient BB 29,292 16,705 Transie.. sc sc

Elements that influence the revenue generation are different type of distribution channel, meal preferences. and customer type.

Here heat map represents that most transient clients book hotels through TA/TO(Travel agency and Tour operator) distribution channel and they tend to prefer BB meals over other types of meals. The high density color depicts the maximum revenue generated by City hotel with customer type "Transient" and distribution channel "TA/TO" along with "BB" as meal type is around 14 million, whereas Resort hotel generated 7 million revenue respectively.

Since most city hotels are centrally located and offer convenient transportation, the most of their guests are "Transients." Transients prefer shorter stays and do not look out additional fun activities.

# Question 3 :

What is the average waiting time for city hotel and resort hotel? 7 6 3 1 January February March April June July 7. 72 August Septem..

October Novemb.. December

052d

The average waiting days for each month over the period of 2014-2017 is visualized using Line Chart. For City hotel, we can notice that average waiting days is approx. 7 whereas 2 days for Resort hotel. City hotel has more demand compared to Resort hotel which indirectly specifies when the demand is more the revenue is more.

# Question 4:

Which hotel is the top choice among the .JtaIy countries? Irélahd' Country • DEU • ESP G3R PRT U •ted Kin 0,937 Ogal Spain 5 ,565 9,610 Morocco 0 2023 Mapbox O OpenStreetMap m Netheilendr Denmark Germany 11,974 Cieéhi; France

18,292 'itzerlånd

The colors in geographical map denotes the countries area, their names along with guest count.

For City hotel, Portugal has the most visitors with 55565 guests, followed by France (18292), Germany (11974), the United Kingdom (10937), and Spain (9610). Similarly when the data is analyzed for Resort hotel, Portugal has the most visitors with 34471 guests, followed by the United Kingdom (13629),Spain (8543),France (3287) and Germany (2224). This clearly depicts guests visiting the City hotel is relatively higher for various countries available in the dataset compared to Resort hotel.