

**REPORT**



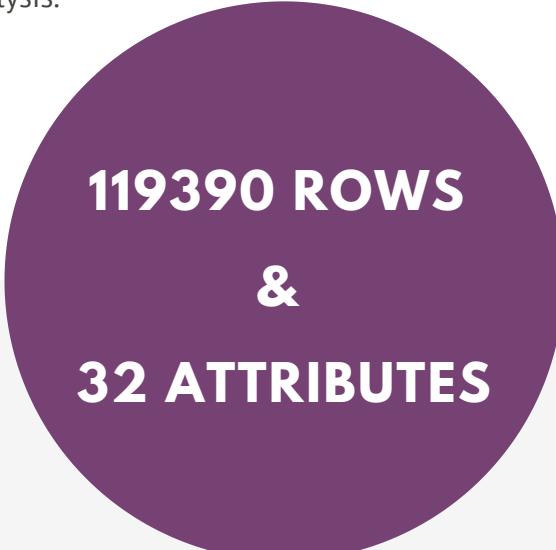
# **ANALYSIS OF HOTEL DEMAND BOOKING DATA REPORT**

**DONE BY:**

ABDULLAH ALHUSSAIN.  
HALAH ALMODARRA  
NOURAH ALMUTLAQ.  
YASMEEN ALDOSSARY

## ABOUT THE DATASET:

Our dataset contains information about the hotels booking such as type of hotel (Resort Hotel or City Hotel), booking time, booking duration, canceled bookings, number of adults, children, etc. The dataset has 119390 rows and 32 attributes varied on numeric and categorical attributes. It has 488 null values in 'country' attribute, 16340 null values in 'agent' attribute, and 112593 null values in 'company' attribute. The original data was acquired from the article Hotel Booking Demand Datasets, written by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019. We choose this dataset as we think that it will be very helpful for people who plan to book a hotel for their next flight and hotel owners who plan to provide offers based on the data analysis.



**119390 ROWS**  
**&**  
**32 ATTRIBUTES**

## DATA SOURCE:

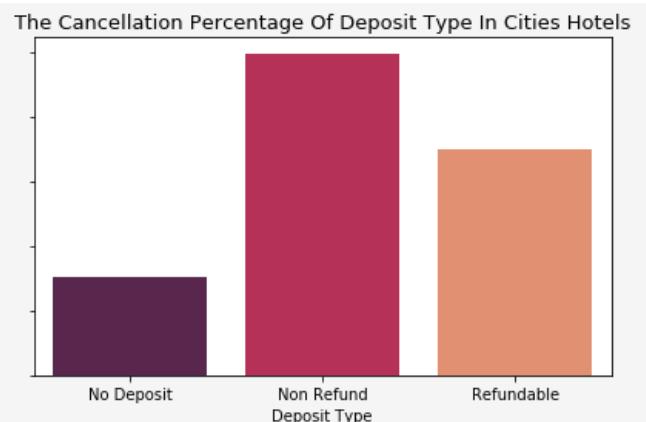
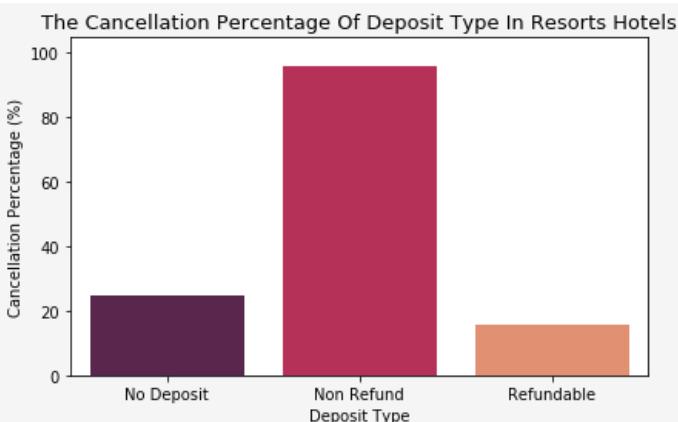
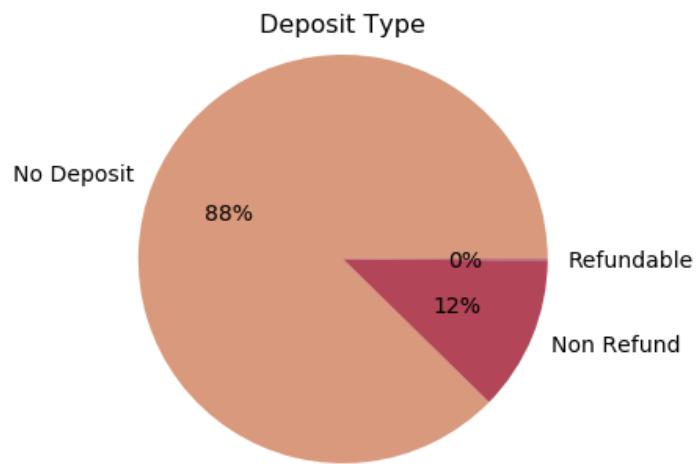
[HTTPS://WWW.KAGGLE.COM/JESSEMOSTIPAK/HOTEL-BOOKING-DEMAND](https://www.kaggle.com/jessemostipak/hotel-booking-demand)

## 1. DOES THE TYPE OF PAYMENT AFFECT THE CANCELLATION OF THE RESERVATION OR NOT?

From the pie chart we can see that 88% of deposit type was No Deposit, 12% Non Refund and 0.14% Refundable. This means that the majority of reservations do not require advance deposit. This leads to our next question, is the No Deposit category the highest category in canceling reservations?

The first plot shows that the highest cancellation was Non Refund followed by No Deposit, then Refundable for Resorts Hotels. And in the second plot we can see that the highest cancellation also was Non Refund followed by Refundable, then No Deposit for Cities Hotels. The results show that customers who pay non-refundable booking fees are more likely to cancel it. We expected that the highest deposit type in cancellation it will be the No Deposit type not the Non Refund type! This is so weird, it made us wonder, was the data collected correctly? Or what happened here?

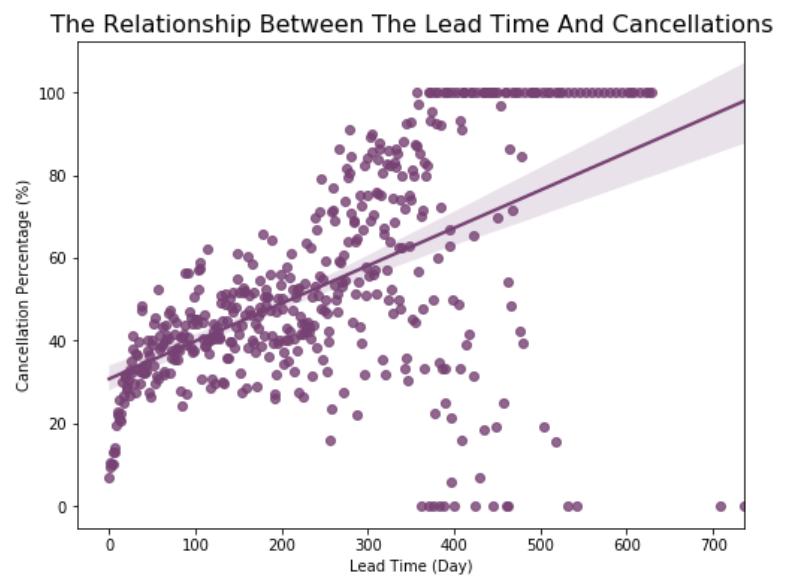
We recommend hotels to make most of their reservations without prepayment, Because it is the highest type in reservations and the least type in cancellations.



## 2. DOES LENGTH OF TIME BETWEEN THE ENTERING DATE OF THE BOOKING AND THE ARRIVAL DATE AFFECT THE CANCELLATION?

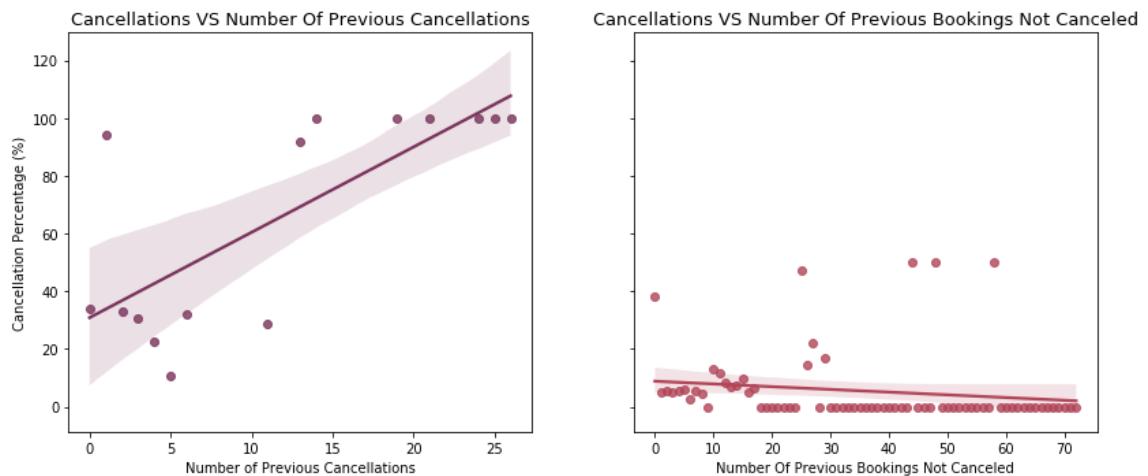
We can see from the plot there's a positive correlation between the cancellation percentage and lead time, that means if the number of days that elapsed between the entering date of the booking and the arrival date increases then the cancellation percentage will increase.

Our recommendation is to reduce the number of days that elapsed between the entering date of the booking and the arrival date as much as possible to minimize and avoid cancellations.



## 3. CUSTOMER BEHAVIOR IN PREVIOUS CANCELLATIONS AND PREVIOUS BOOKINGS NOT CANCELED

In the first plot there's a positive correlation between the cancellation percentage and the number of previous cancellations, that means if the number of previous cancellations increases then the cancellation percentage will increase. And in the second plot there's also a correlation



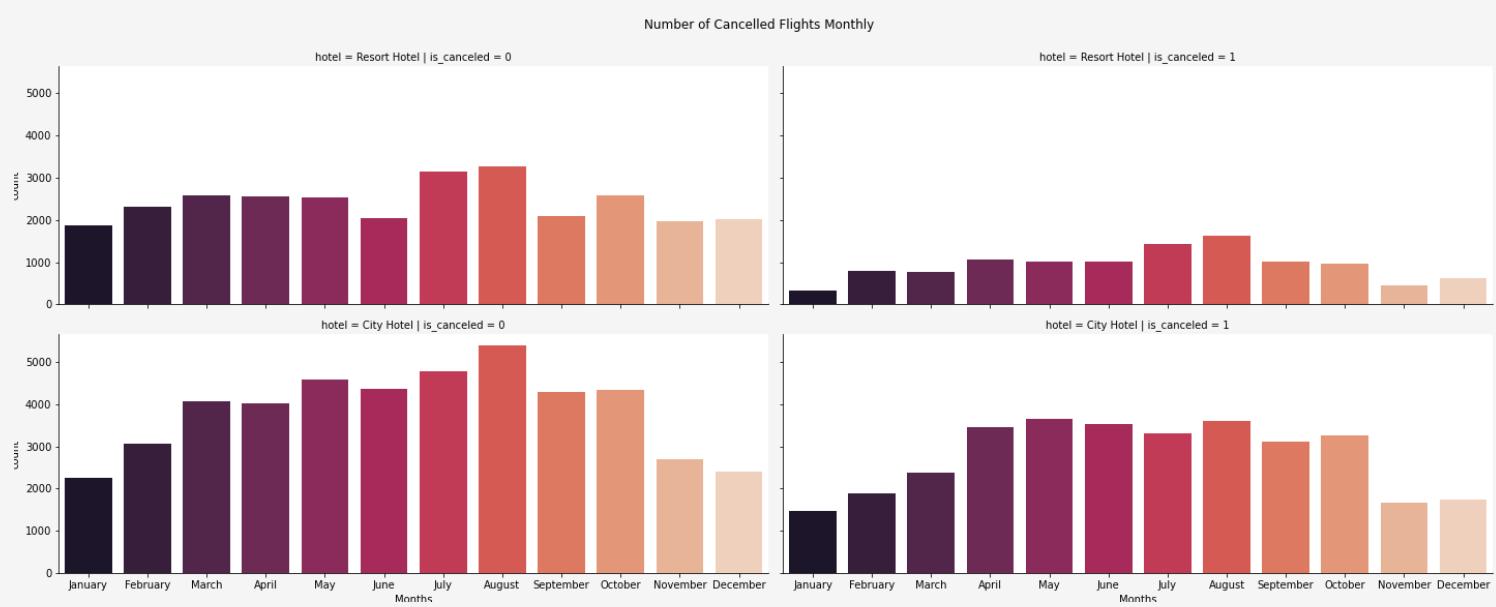
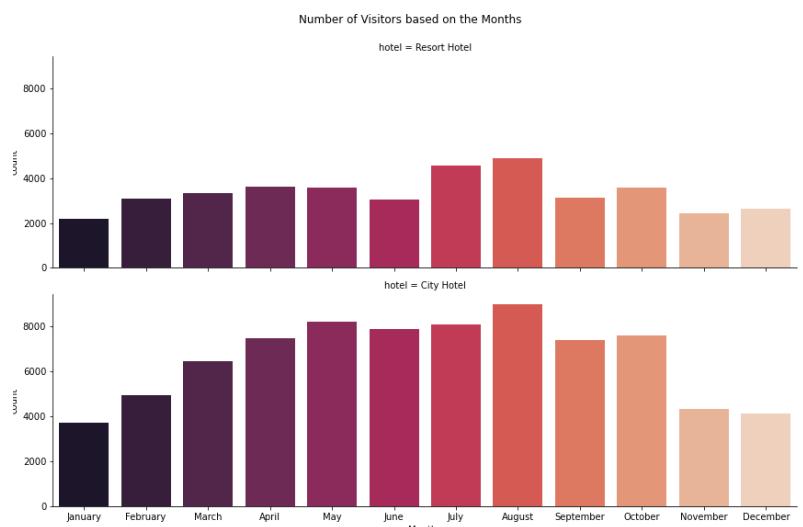
between the cancellation percentage and the number of previous bookings not canceled but it's a negative correlation, that means if the number of previous bookings not canceled increases then the cancellation percentage will decrease.

In other words we can say that the customer who has a large number of canceled reservations tends to cancel the upcoming one, and the customer who has a large number of reservations tends to keep the reservation.

We recommend hotels to increase retention rate of their loyal customers, And put a limit for the number of cancellations per customer to avoid customers who cancel reservations frequently, And send a survey to find out what's the reason of the frequent cancellations and how they can improve their service.

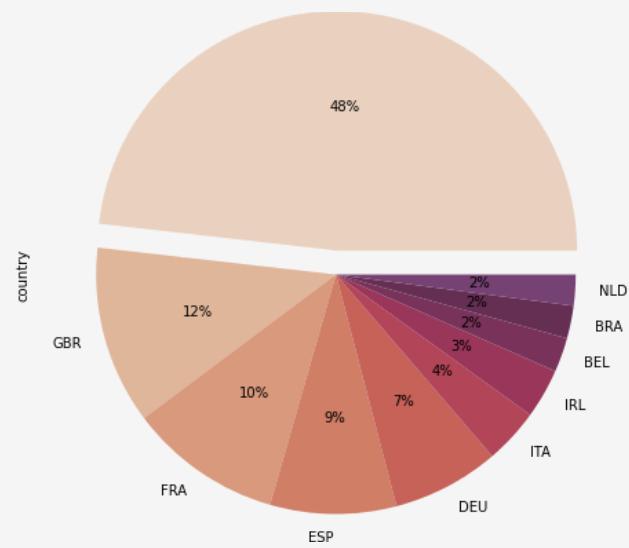
## 4. WHICH MONTHS HAVE THE LOWEST NUMBER OF VISITORS?

In the first chart, we aim to find the months that have the lowest number of visitors. As we can see November, December and January have the lowest number of visitors. This leads us to go for more investigation to see if the lowest number of visitors due to the increase in the number of the canceled flights. So, in the next graph, we visualized the same graph but with the number of canceled flights. However, the number of canceled flights was the lowest in the same three months: November, December, and January. Therefore, we would encourage hotels and airlines to increase the promotions on prices of the tickets and hotel booking in these three months to increase the revenue.



## 5. WHICH TOP 10 COUNTRIES THAT VISITORS COME FROM?

For the second insight, we display the top ten countries that visitors come from which are: The Portuguese Republic (PRT), United Kingdom of Great Britain and Northern Ireland (GBR), France (FRA), Spain (ESP), Germany (DEU), Italy (ITA), Ireland (IRL), Belgium (BEL), Brazil (BRA), and Netherlands (NLD). Therefore, this insight might help hotels and airlines to provide special packages to these countries. Also, this might help governments to facilitate the issuing of the visa to attract more visitors which will be reflected in the country's economy.



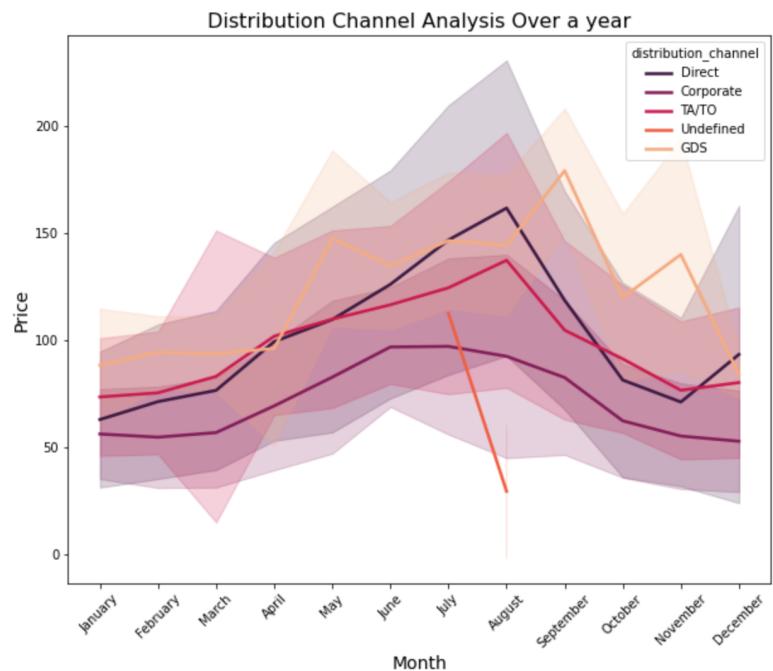
## 6- BOOKING PRICE BY CUSTOMER TYPE



The given boxplot chart illustrates the booking price by customer type over the three years. There are four customer types which are contract, group, transient, and transient party. As the chart showed that the highest cost of room booking was booked by a transient. The Transient is the booking booked by an individual that is not part of a group or contract. Thus, targeting this group of customers by preparing marketing strategy will increase the sales.

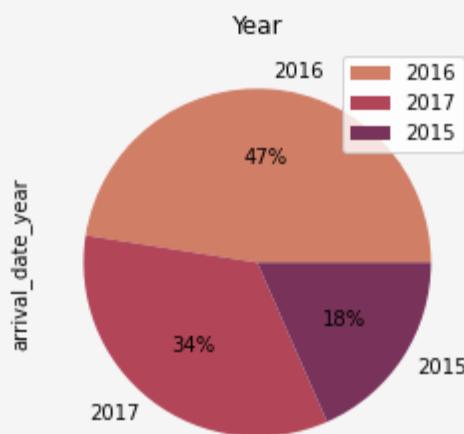
## 7- DISTRIBUTION CHANNEL ANALYSIS OVER A YEAR

The following chart present the analysis of booking distribution channels within a year. The distribution channels are the way that booking have been done. There are four TA/TO, direct, corporate, and GDS. TA/TO is Travel Agents and Tour Operators. The direct booking is the highest booking number in August which is the summer vacation. Moreover, the heist price reaches the price of around 180 in September booked by GDS. Thus, targeting this channels during the summer will rise the booking number.

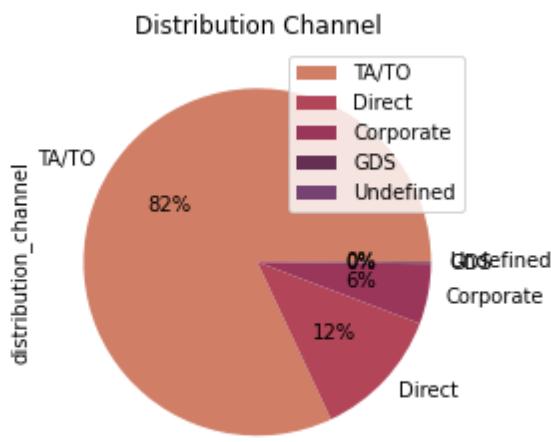


## 8. ANALYSIS OF BOOKING WITHIN 3 YEARS

This plot shows the number of customers arriving at the hotels each year. The data set contains records for three years which are 2015, 2016 and 2017. It is obvious from the plot that year 2016 is the highest year in terms of the number of customers. Year 2017 is the second highest and year 2015 has the least number of customers among the three years.



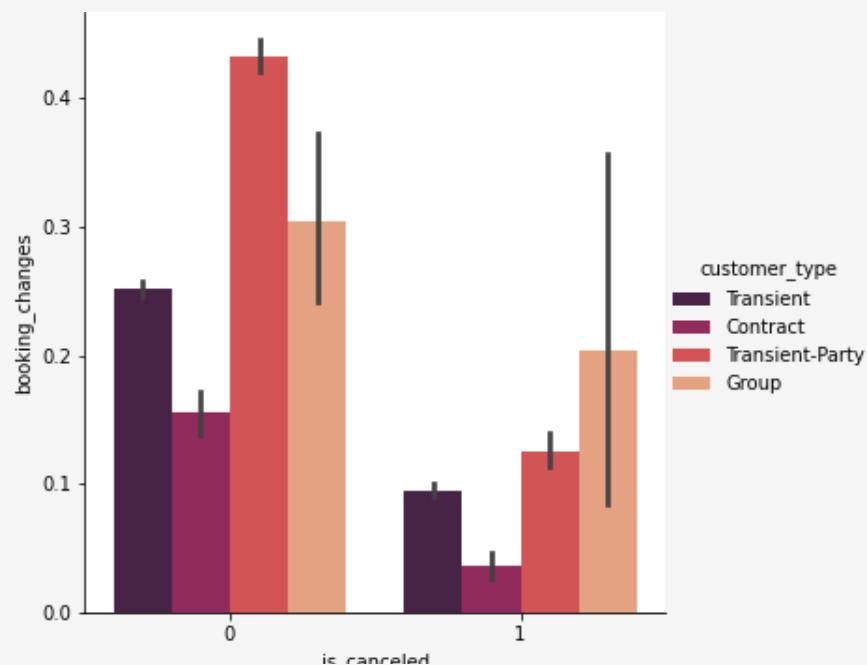
## 9. PERCENTAGE OF DISTRIBUTION CHANNEL



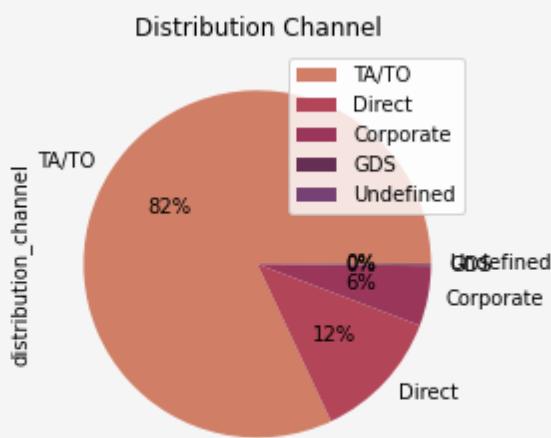
This plot shows how frequently each distribution channel is used for booking rooms in the hotels. It can be clearly seen that most customers prefer TA/TO in the first place since more than 80% of them reserve their rooms through this channel. Secondly, 12% of the people prefer direct reservation as a channel for booking a room. The rest of the people prefer other distribution channels such as corporate and GDS.

## 10. RELATION BETWEEN BOOKING CHANGES AND CANCELLING THE RESERVATION

We created a plot that shows the relationship between number of times of changing the reservation and number of reservation cancellations. Results show that people who change the reservation more often are less likely to cancel their reservations. This is logical because changing the reservation time means the customer is satisfied with the hotel but the issue is with the timing only, so, if they change the reservation time, they are less likely to cancel it.



## 11. NUMBER OF PEOPLE CANCELLING RESERVATION IN EACH HOTEL



For this insight, a plot has been created showing the number of reservation cancellations for each type of hotel, i.e., city hotels and resort hotels. Results show that almost half of the people who book a room in city hotels cancel their reservation, and almost a quarter of people who book a room in resort hotels cancel their reservation. For city hotels, this requires further investigations about the reasons that make half of the people cancel their reservations.

