

# Hotel Reservation Data Analysis Executive Presentation



Date: April 3, 2020

Presenter: Justin Albright

# Can hospitality managers predict guest cancellations in order to optimize revenue?

---

!

Cancellations are a major part of any hospitality business that accepts reservations. Every cancellation is lost revenue, so managers are encouraged to find ways to mitigate this.

Reservation data from hotels can be analyzed to answer this question. We can then develop a framework to hedge against future lost revenue from cancellations.

!

With so many unique geographies and hospitality experiences, it can be challenging to identify where to start and what to analyze. Much of this information is proprietary and confidential for business purposes.

Detailed information from nearly 120,000 reservations is available from two hotels in Portugal spanning over two years (July 2015 – August 2017). 32 distinct pieces of data were collected from each reservation, some numeric and some categorical. Names and other personally identifying information were removed to ensure rights to privacy.

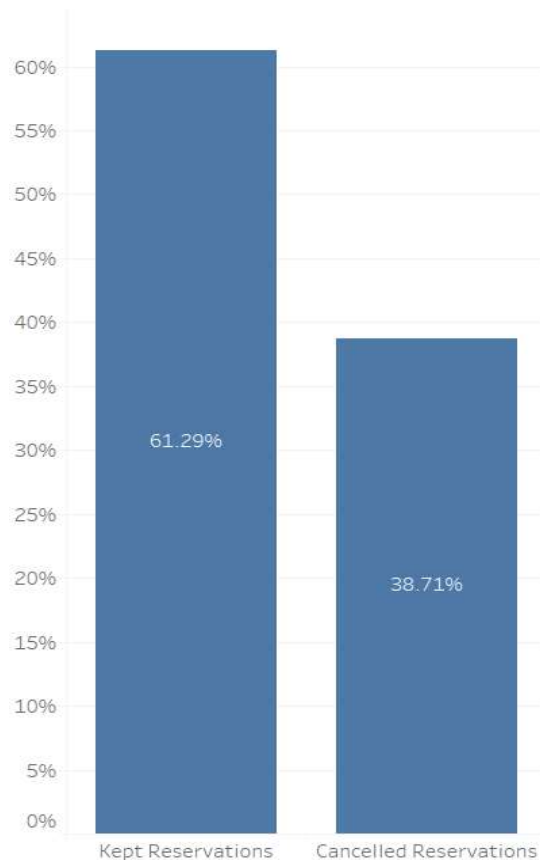
!

Specifically what did we analyze?

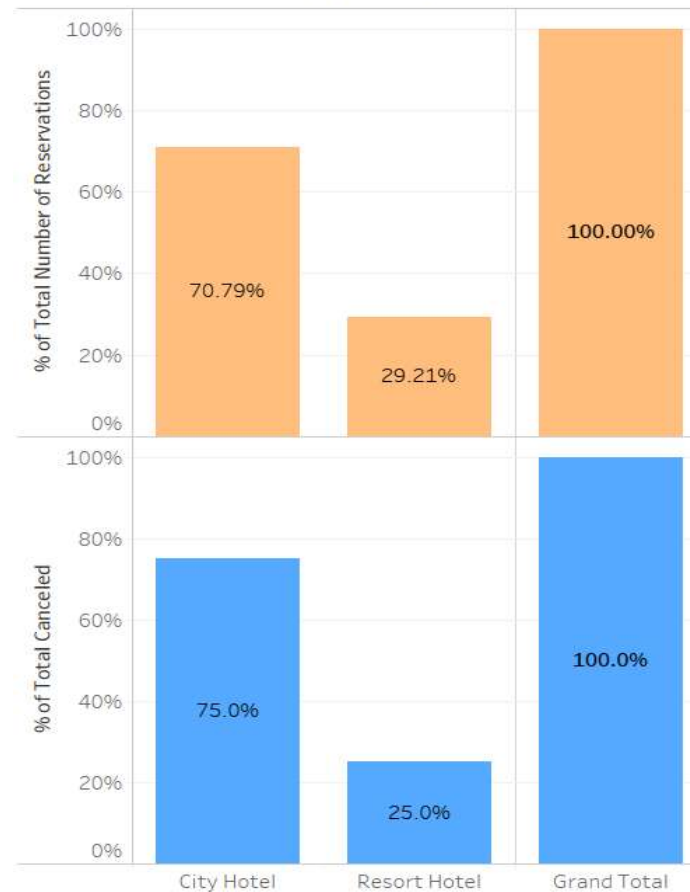
We analyzed the data of guests who cancelled vs. those who didn't in relation to the other 31 variables (pieces of information) documented under each reservation. Statistical analysis of this large data set was performed in Python, with further visualizations prepared in Tableau. This presentation reports on what patterns emerged and what customer types are most likely to cancel.

## Cancellations Overview and Hotel Types

Total Reservations - Kept vs. Cancelled



Reservations by Hotel and Proportion of Cancellations

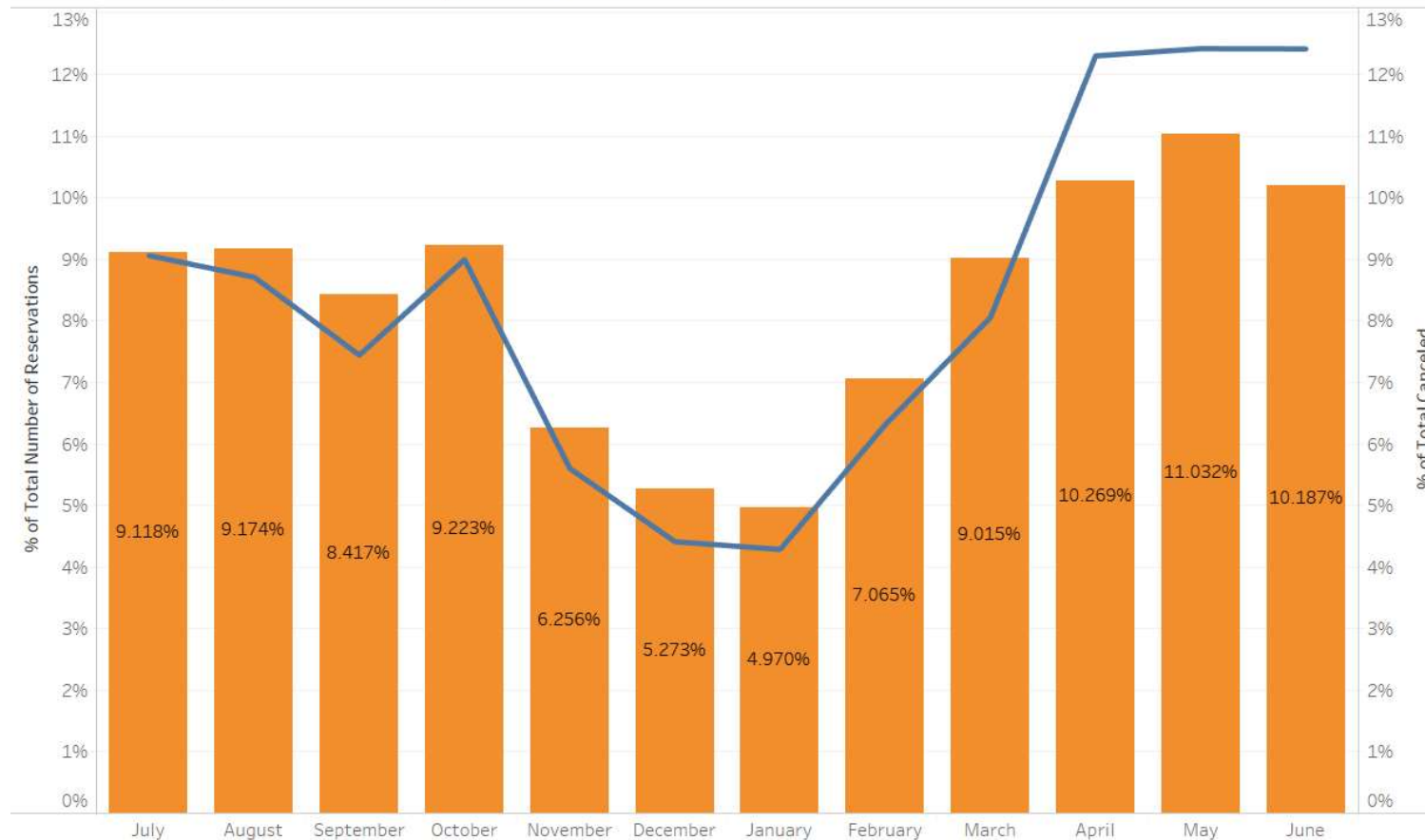


### Insights

- Over one-third of all guests cancelled their reservation
- Over 2/3 of all reservations were made at the City Hotel vs. the Resort Hotel, although a higher proportion of cancellations occurred at the City Hotel (3/4 of total).

## Seasonality

Reservations by Month and Proportion of Cancellations

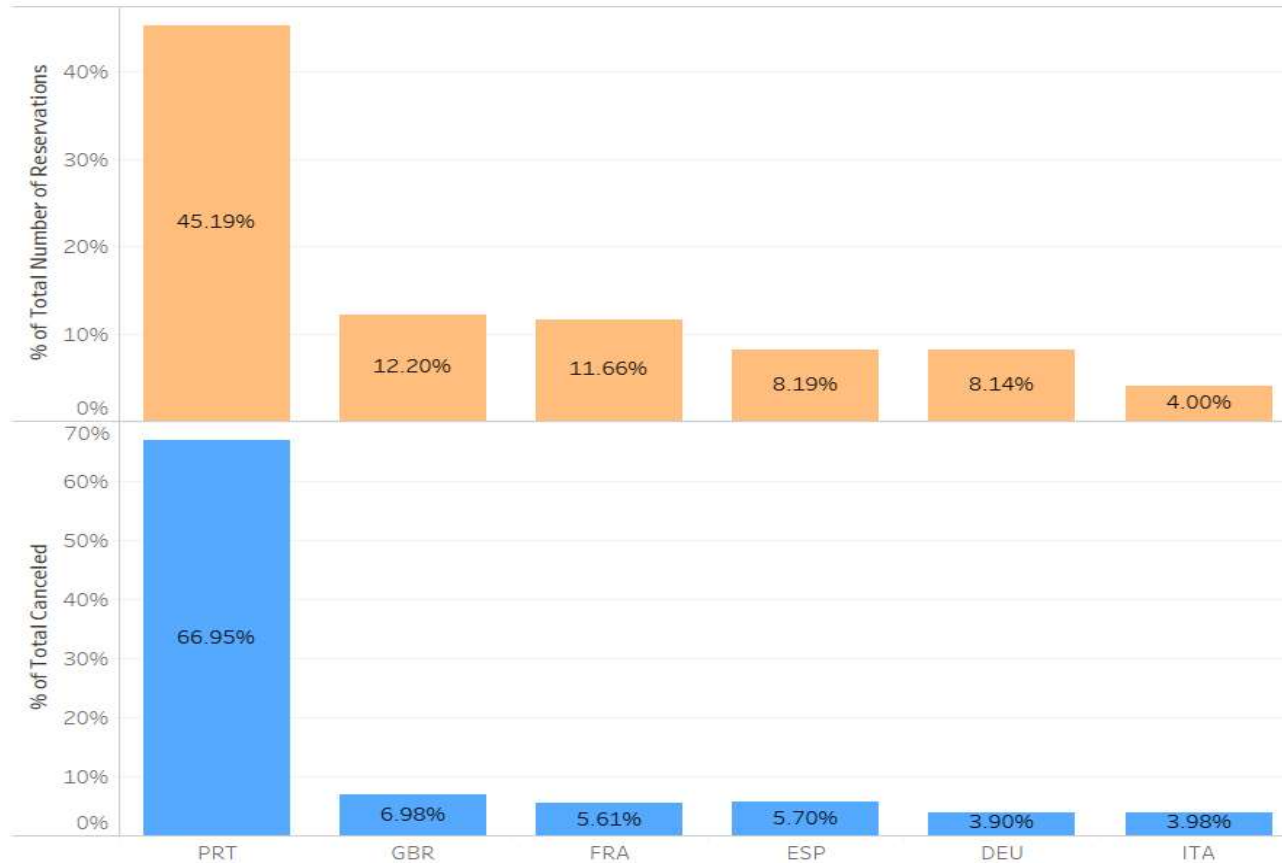


### Insights

- **Reservations follow a seasonal pattern common in Portugal.**
- **Cancellations follow little discernible pattern itself and overall proportionately a product of the number of total reservations made. (Slightly higher April-June)**

## Guests' Country of Origin

Reservations by Nationality and Proportion of Cancellations

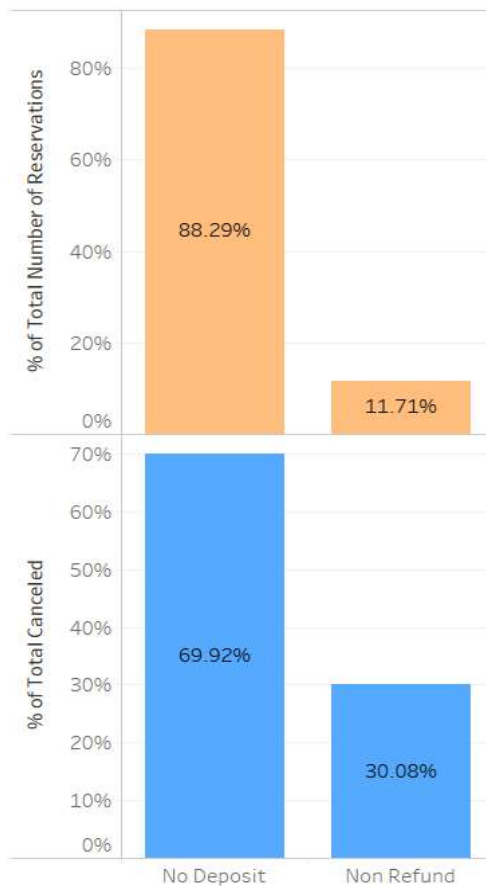


### Insights

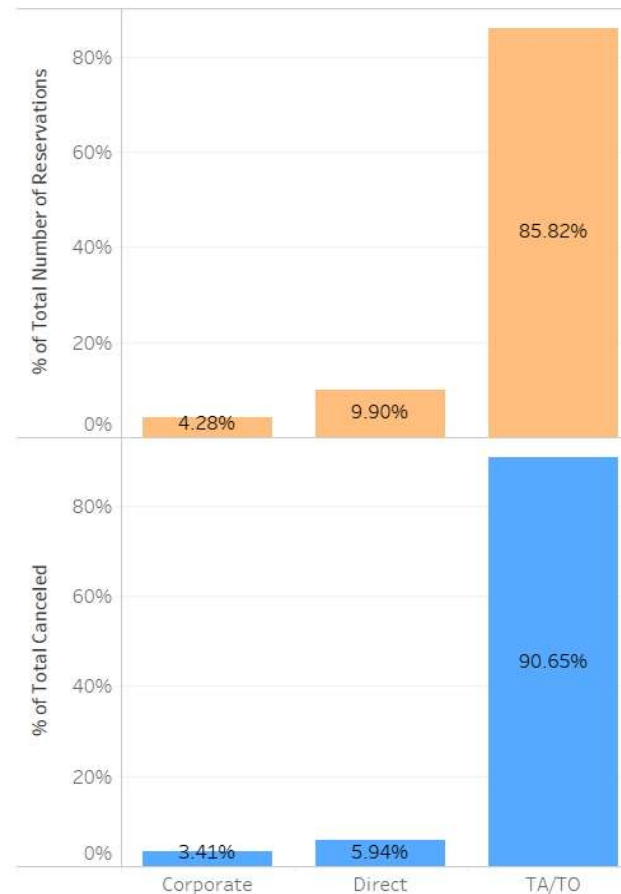
- Almost half of all guests are local to Portugal, yet they make up a large majority of cancellations (2/3).
- Over 90% of guests are from Europe. Every other nationality contributes a disproportionately small share of cancellations.
- Note – guests come from over 150 countries, but only top few shown here.

## Distributions and Deposits

Reservations by Deposit Type



Reservations by Distribution Channel

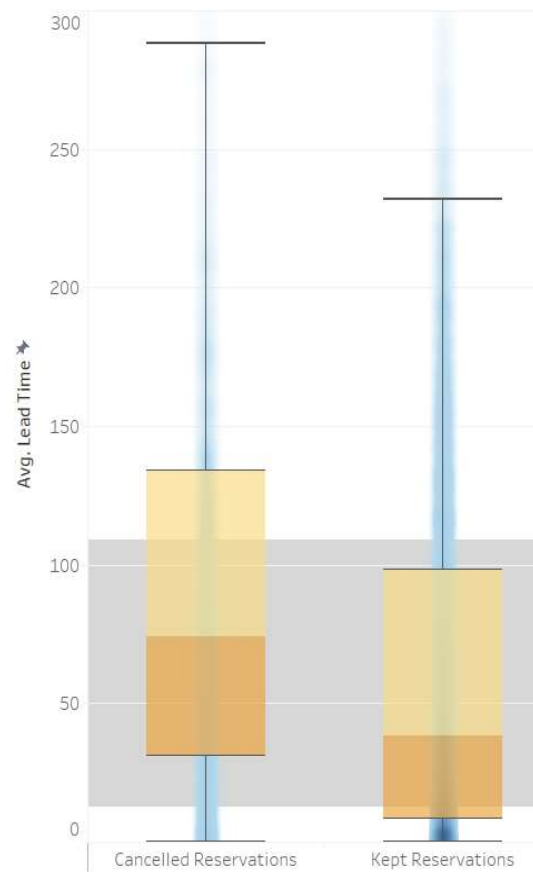


### Insights

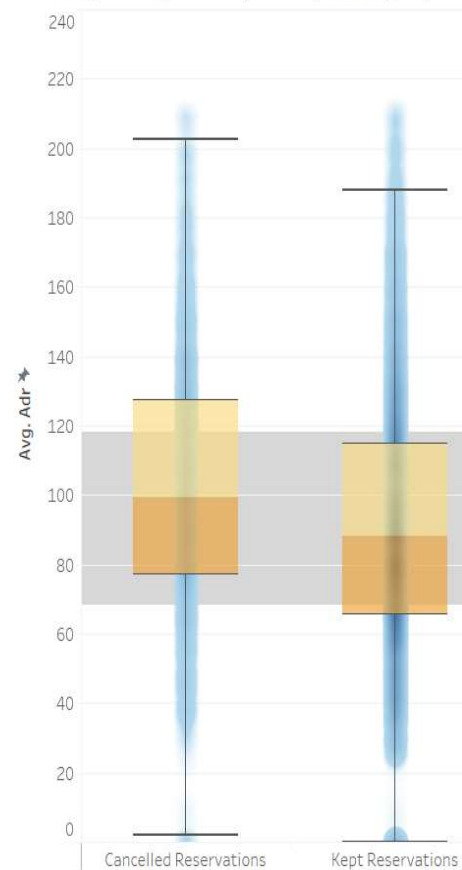
- Counterintuitively, those who put down a deposit are more likely to cancel. Less than 1/8 of guests placed a deposit, but they account for almost 1/3 of cancellations.
- Most guests (over 4/5) used an agent or operator to reserve, which is common in Europe. This majority was slightly more likely to cancel than those who booked directly or went through their company.

## Boxplot Analysis of Other Variables

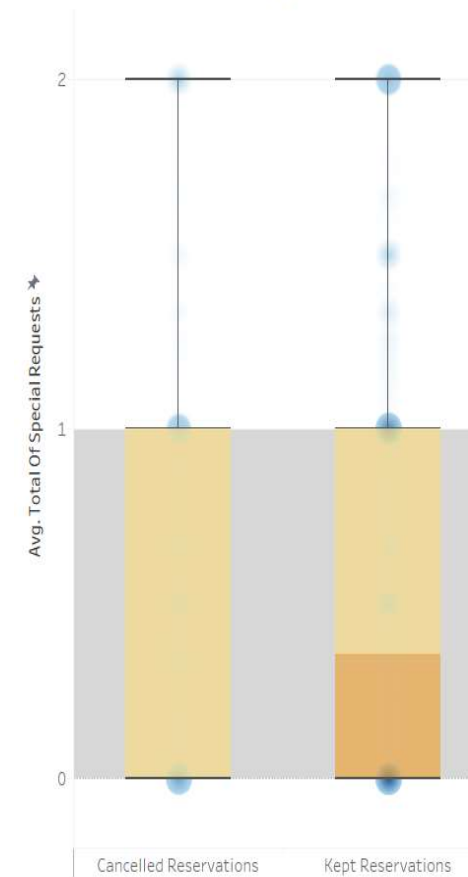
Lead Time in Days



Average Daily Rate (Price per Night)



Number of Special Requests by Guest

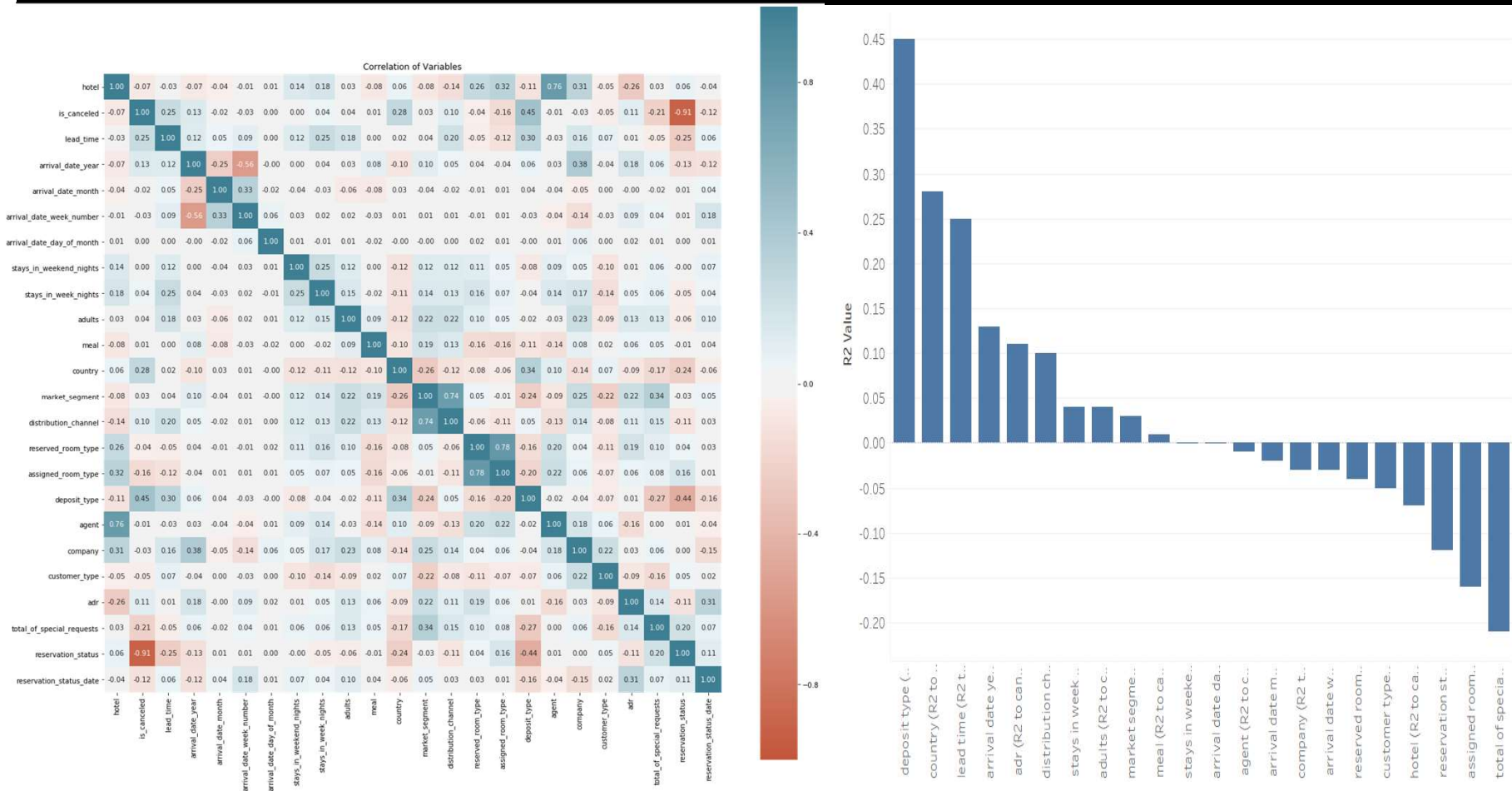


### Insights

1. Lead time – longer lead times result in a higher cancellation rates.
2. ADR – higher-priced rooms drive a higher cancellation rate.
3. Special requests – more requests are made by those who don't cancel. Those who keep their reservation are more likely to invest special effort ahead of time for their stay.



# Relationships between All Variables



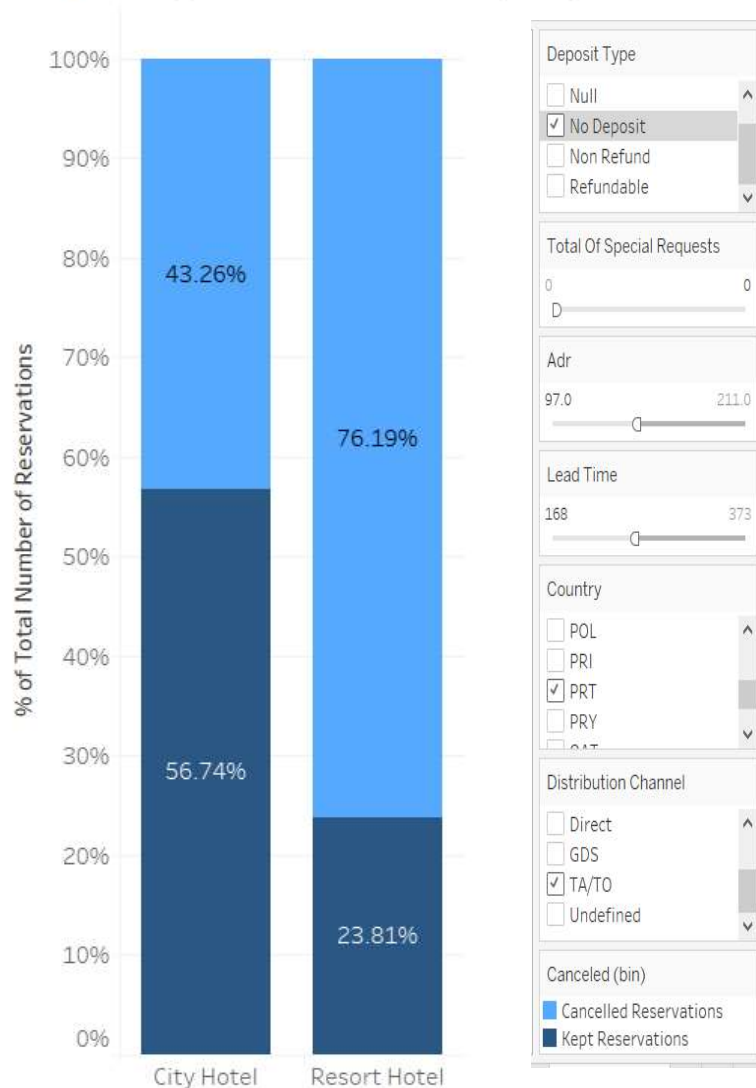
## Insights

- Matrix (left) – the darker the color, the more useful the information is to predicting cancellation.
- Bar plot (right) of variables is in descending order relative to its usefulness in predicting cancellation. Deposit type, country (nationality), and lead time, the first 3 variables here, is the information most likely associated w/cancellations.
- This is supported by the relationships shown in the previous slides.



## Application Example – Determining the Real Odds of Cancellation

Probability of Cancellation by Key Factors



### Insights

- Using the 7 most correlated pieces of information submitted by each guest determines the probability of cancellation.

#### Example shown:

- Portuguese guest
- Price of 97 euro per night
- No deposit
- Lead time of 168 days
- Booked through an agent (distribution channel)
- This guest has a 43% chance of cancelling at a City Hotel.
- The guest has a 76% chance of cancelling at a Resort Hotel.

## Conclusions – Practical Use for the Hospitality Manager

---

**Using guest profiles as such, cancellation odds for each guest are predictable. Taken altogether, determining cancellation probability for an entire hotel for future dates becomes feasible.**

### **Example:**

- 80 guests have booked rooms at the city hotel
- Guests' combined cancellation probability is 40%.
- Hotel manager can anticipate that day only 48 guests arriving:  $((80 - (40\% \times 80)) = 48)$ .

### **Other uses to enhance hotel performance:**

- Accurate revenue forecasting.
- Overbooking to maximize revenue. Especially during high season a hotel can confidently overbook to a certain point, based on cancellation odds, to maintain full occupancy.
- Operations - scale staffing, services, and supplies based on the anticipated number of guests.
- Track hotel's outreach to potential guests (i.e. deposits, special requests, how far in advance bookings are accepted).