



FINAL PROJECT – HOTEL BOOKINGS

YASH PISE

Table of Contents

1	Final assignment – Hotel Booking Data.....	2
1.1	How to reduce cancellation of HOrl Rooms.....	Error! Bookmark not defined.
1.1.1	Introduction	2
1.1.2	Data Preparation	3
1.1.3	Exploratory Data Analysis.....	5
1.1.4	Summary	12
1.1.5	Formatting & Other Requirements	Error! Bookmark not defined.

Table of Figures

Figure 1	Carvana's Car Vending Machine in Tennessee.....	Error! Bookmark not defined.
Figure 2	Carvana Stock Price from IPO to Current.....	Error! Bookmark not defined.
Figure 3	Carvana's Slowing Revenue Growth.	Error! Bookmark not defined.
Figure 4	Carvana's Increase in Spending.	Error! Bookmark not defined.
Figure 5	Carvana's Existing Markets Account for 73% of Population.....	Error! Bookmark not defined.
Figure 6	Auction Companies vs. Bad Buys Comparison Chart.	Error! Bookmark not defined.
Figure 7	Dashboard of Bad Buy Cars by Nationality.	Error! Bookmark not defined.

Table of Tables

NO TABLE OF FIGURES ENTRIES FOUND.

FINAL ASSIGNMENT – HOTEL BOOKING DATA

Introduction

Problem statement

Optimization of occupancy has a big impact on profitability, and cancellation has a big impact as well. Owing to the prospect customer's inability to reserve a room due to availability, booking cancellation results in a significant loss for hotels. Therefore, the primary issue a hotel faces is how to lower cancellations

Plan of Implementation

We will thoroughly examine the cancellation data. In order to prevent cancellations or at least predict them, try to spot trends in them and respond accordingly on those causes.

Analytic technique

Start with the source of the booking, we will move on to several aspects that are important in cancellations, examine trends from prior data, and finally, forecast the outcome so that we can have an idea about it and take proactive steps to reduce cancellation. We'll use various visualizations to support our debate and conclusion.

Outcome of Analytics

It will assist Hotels in determining things like which Agents to penalize, which nations to have strong cancellation policies, how to give repeat customers discounts, and how to maximize occupancy by anticipating cancellation

Data Preparation

Data Source

Data Source of this project has been taken from

<https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-02-11/readme.md>

The data this week comes from an open hotel booking demand dataset

Data Variables

Data set have 32 variables. The data this week comes from an open hotel booking demand dataset

The data consists of 119390 rows and 32 columns

Variables are: -

```
[1] "hotel", [2] "is_canceled",
[3] "lead_time", [4] "arrival_date_year",
[5] "arrival_date_month", [6] "arrival_date_week_number",
[7] "arrival_date_day_of_month", [8] "stays_in_weekend_nights",
[9] "stays_in_week_nights" [10]"adults",
[11] "children", [12]"babies",[13] "meal",
[14] "country", [15] "market_segment", [16] "distribution_channel",
[17] "is_repeated_guest", [18] "previous_cancellations",
```

[19] "previous_bookings_not_canceled", [20] "reserved_room_type",

[21] "assigned_room_type", [22] "booking_changes",

[23] "deposit_type", [24] "agent", [25] "company",

[26] "days_in_waiting_list", [27] "customer_type", [28] "adr",

[29] "required_car_parking_spaces", [30] "total_of_special_requests",

[31] "reservation_status", [32] "reservation_status_date"

Data Cleaning

In order to better comprehend the data and extract findings to develop a strategy, it is essential to clean the dataset in accordance with our research goal. I had to eliminate certain variables from the Hotel data set because they did not support the objective. Those that received a different classification, null values, and so forth.

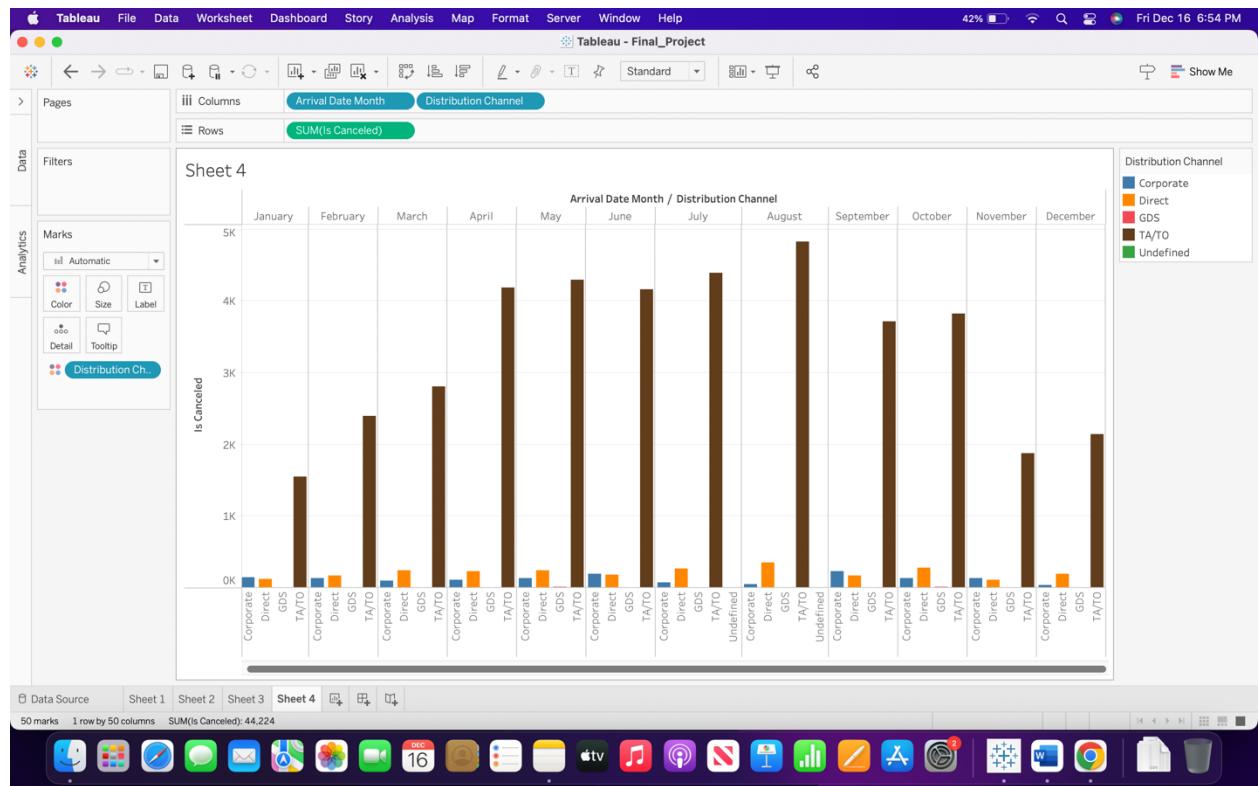
Methods /Tools for data Cleaning:

Data Cleaning: Excel

Data Interpretation: Tableau

Exploratory Data Analysis

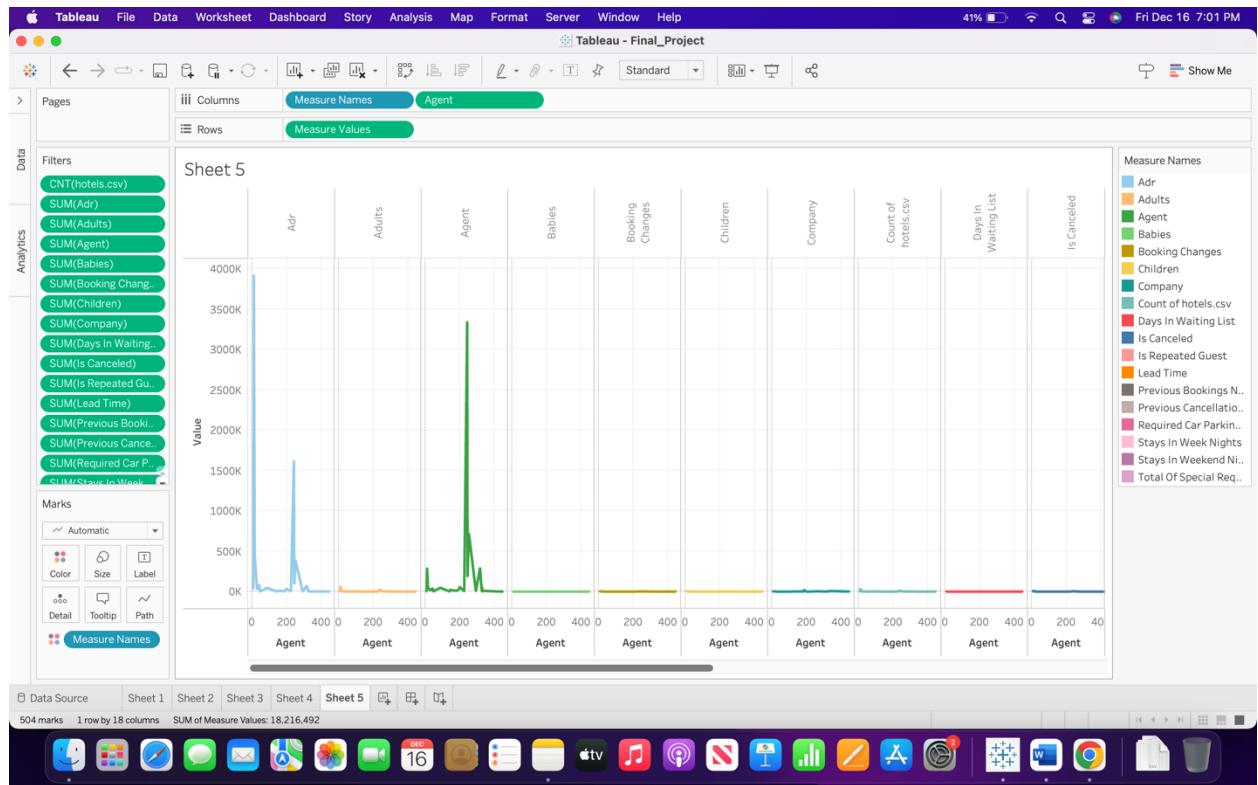
1. Distribution channel/ Arrival date VS Cancellation



With internet booking being so simple, the distribution channel also plays a crucial role. As a result, we should pay alternative outlets with lower cancellation rates a higher or equivalent commission.

Here, I have analyzed the cancellation of customers with regards to distribution channel. We can see in the above graph that the month of August has the most cancellation among all in all the groups. TA/TO has the highest number of cancellations whereas GDS has the lowest.

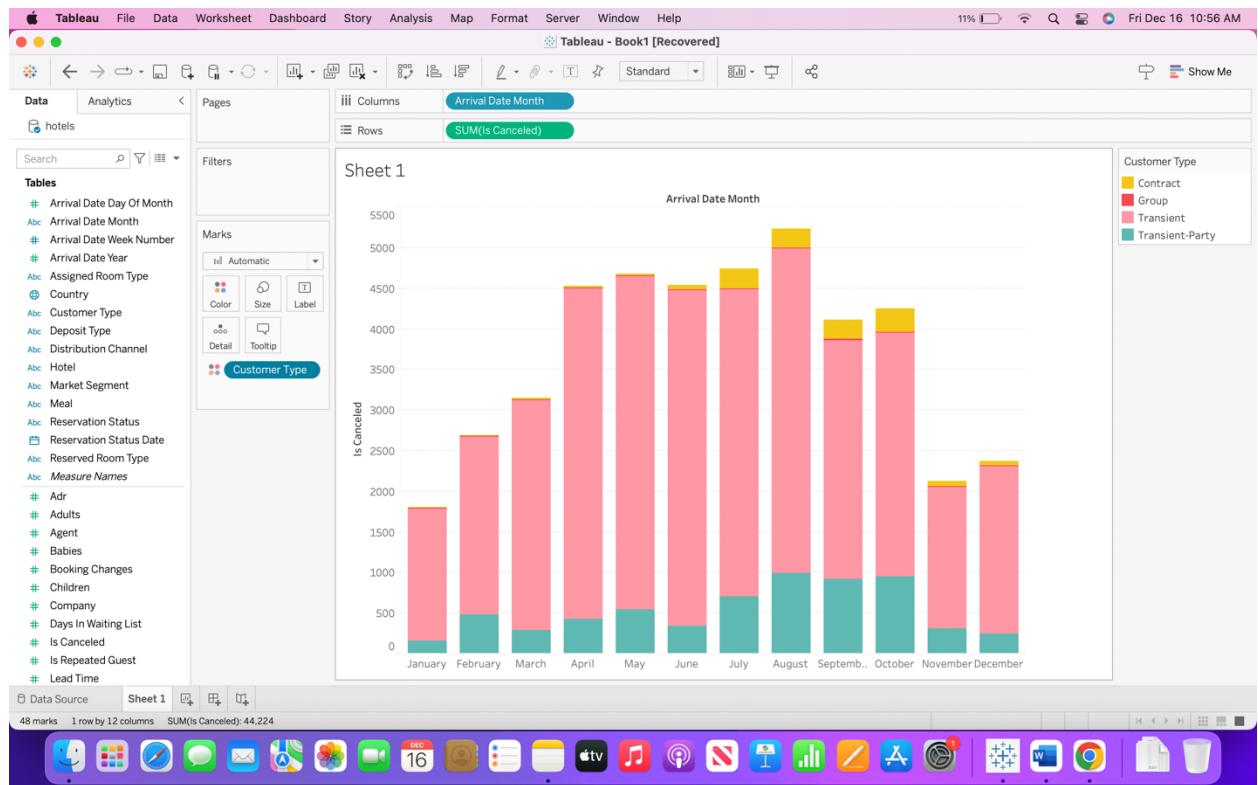
2. Value v/s Measure Names



Value of measure names is very crucial here as we can target the correct audience to market the hotels using digital marketing and other number of sources.

Here we can see the comparison between measure names and value. Adr has the highest value followed by Agent and Lead time. Adults, babies, booking changes, children, company, etc has the lowest value.

3. Customer Type v/s cancellation

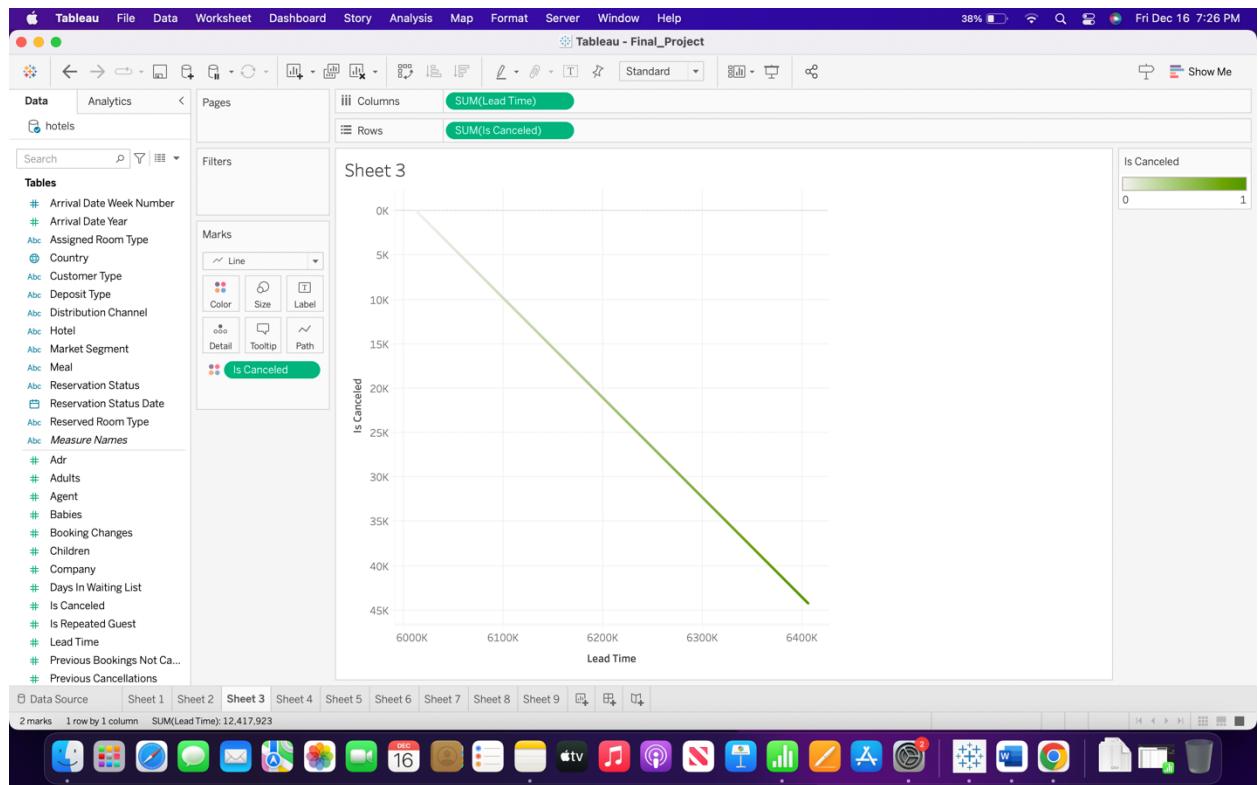


We can infer that transient customers will cancel more frequently, so discounts will be granted in accordance with the kind of booking and can be subject to cancellation fees. The group type with the fewest cancellations can receive free cancellation from us.

Here, we can clearly see that transient customers are the one with the highest number of bookings in each month whereas group customers type has the lowest.

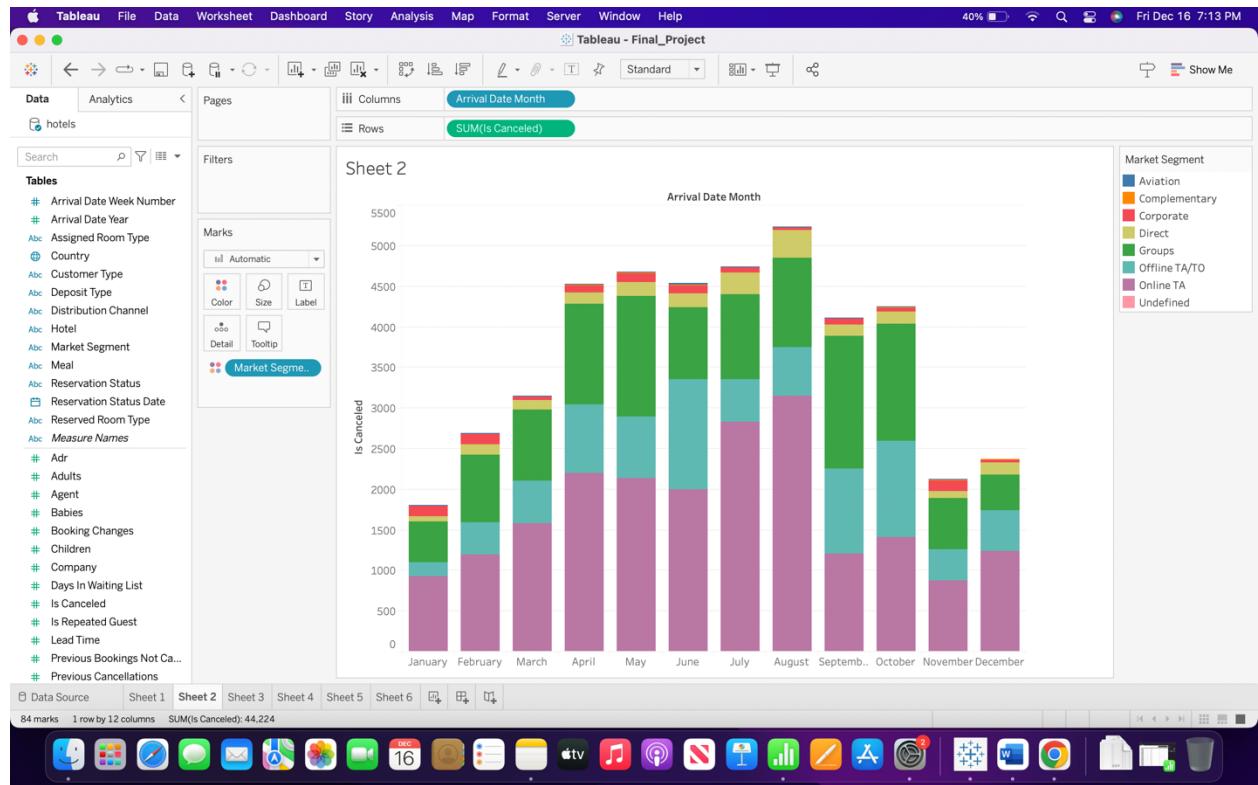
This will help in targeting the audience to market about the hotels.

4. Lead time v/s Cancellation



We can clearly see here that as the lead time increases the cancellation rate decreases. We can say that we should always try to keep the lead time more to reduce the cancellations.

5. Cancellations according to market segments

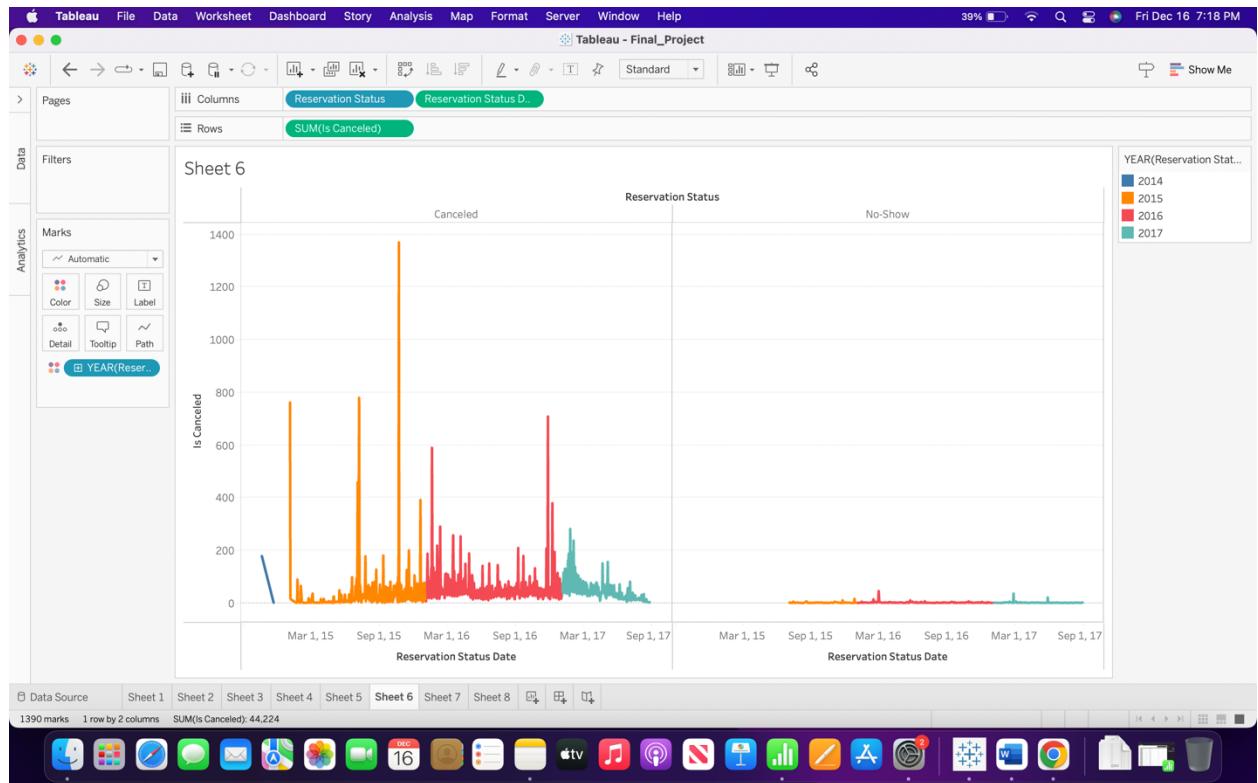


Due to the convenience of booking online, distribution channels are also quite significant. As a result, we should provide more or, in this case, equal commission to channels that see fewer cancellations.

Here, Online TA market segment is the one with highest number of cancellations followed by group, and Offline TA/TO. To avoid these cancellations these market segment customers should be treated by extra hospitality and should be given some more rewards or points so that they are happy.

Aviation market segment customers seems happier with the hotels as they have the lowest number of cancellations followed by corporate market segment.

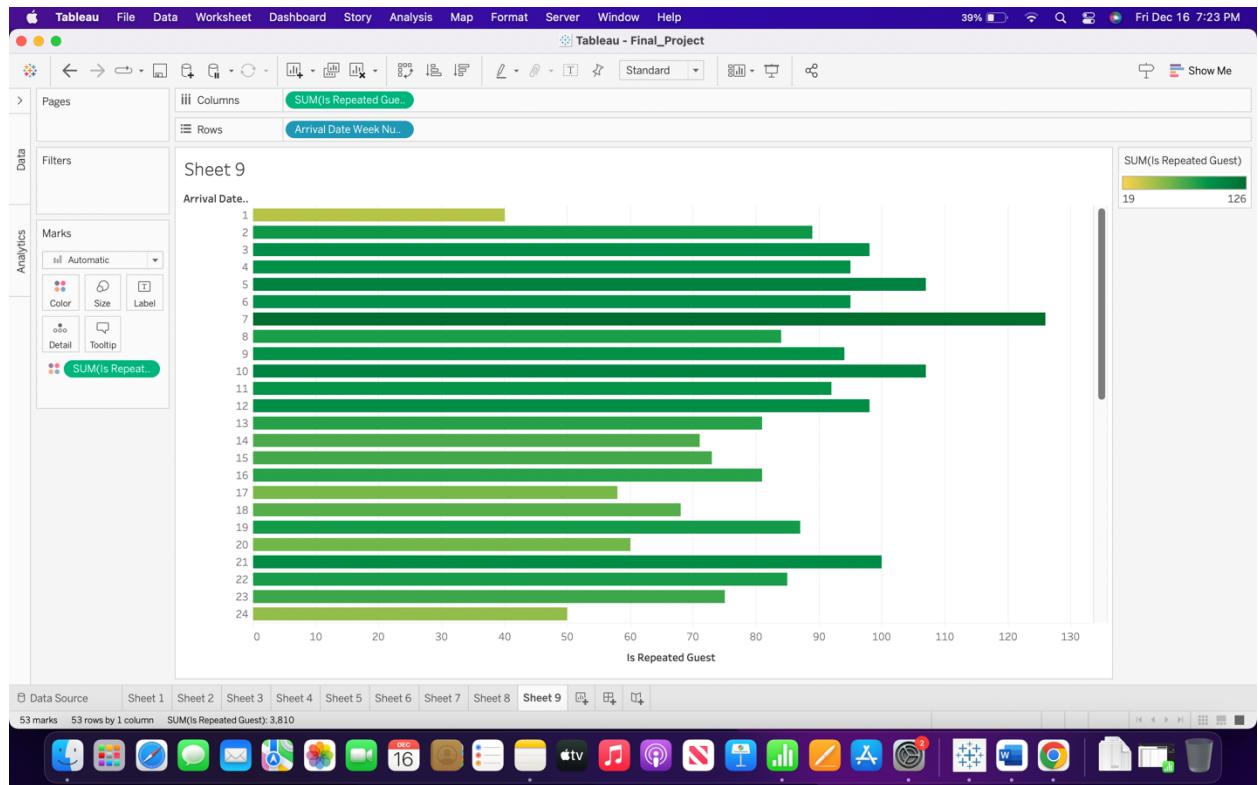
6. Reservation Status



Here we can see that most of the time customers cancel the reservation. Also, this makes sense as cancellation gives them at least some of the refund than doing a No-Show. This graph explains the reservation status between years 2014 – 2017. Year 2015 is the one highest numbers of cancellations whereas year 2017 is one with the lowest numbers of cancellation.

On the other side if you see closely, 2016 is the year with highest numbers of No-shows and 2015 is with the lowest. However, the number hasn't changed drastically so we can say that the graph of no-shows does not varies more.

7. Number of repeated guests each week



Here I have analyzed the number of repeated guests weekly. I have amazed by the number of repeated guests each week. This clearly shows that the customers are loyal to the hotels.

This also tells that customer services of hotels is awesome which in turn gives them repeated customers.

Also, week 7 is the one with highest number of repeated guests followed by weeks 5, 10, and 22.

Week 1 has the lowest number of repeated customers followed by week 24.

Summary

We have examined numerous factors that may affect cancellation in depth and have drawn valuable conclusions from them. In summary, the last quarter had a significant number of cancellations, most of which occurred in the month of August. Booking with greater lead time is advantageous as we have noticed very few cancellations in them. Online cancellation is highly high, thus need to set strong policy on that. A partial deposit should be collected in advance to discourage cancellations, and City Hotel should solicit feedback to determine why customers are cancelling so frequently.

With the assistance of this, we can assist hotels in anticipating when they need ask customers for partial payments in order to prevent cancellations and when they can waive them off. The investigation is limited by the lack of customer feedback data. Additionally, if we had access to competitive hotel data, we could have compared them and concluded about where we stand in terms of market average cancellation rates and which variables are shared by all. It would be very helpful to corroborate the variables that we identified if feedback data was supplied.