

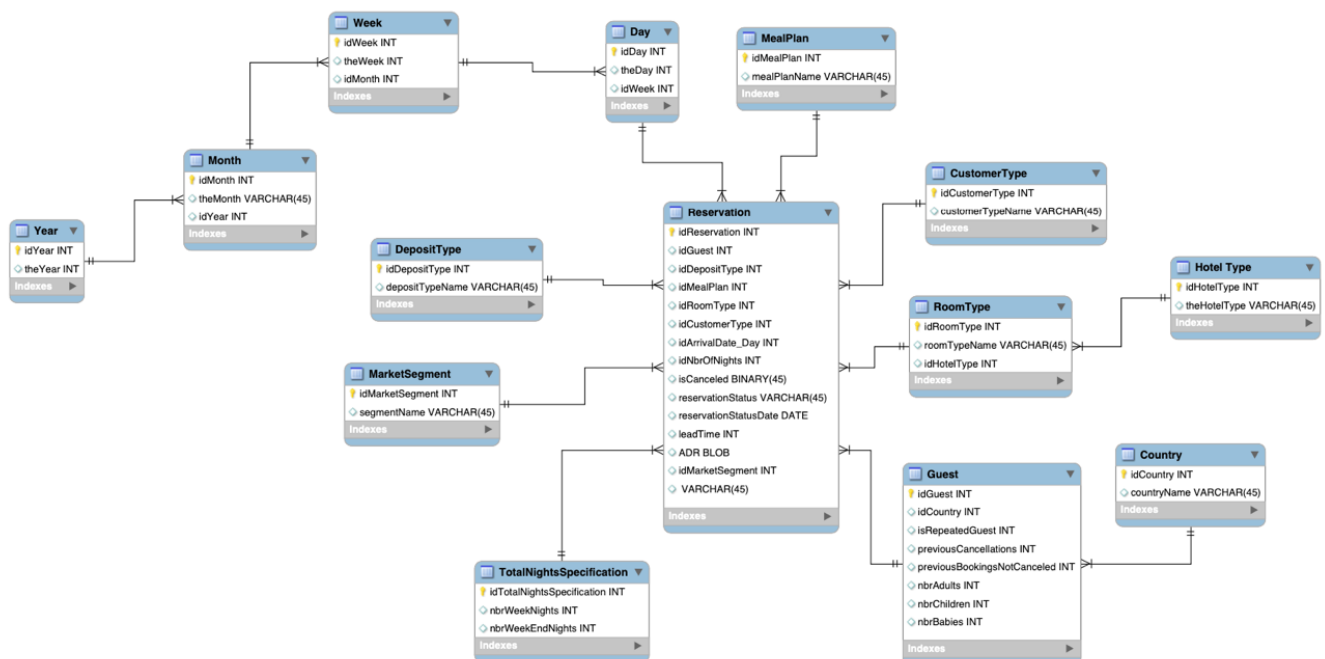
Data warehouse Report: Step 1 2

1. **Data Set:**

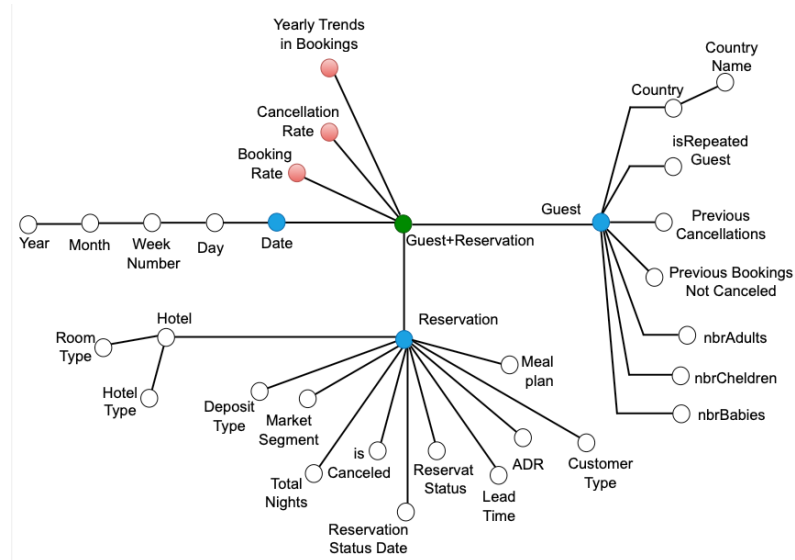
I have selected a **Hotel Booking dataset** from Kaggle. The dataset I've chosen provides comprehensive information on hotel reservations, encompassing 119,390 observations for both a City and a Resort Hotel, including both arrivals and cancellations. For detailed information about the columns and their descriptions, please refer to the Kaggle link provided:

<https://www.kaggle.com/datasets/khairullahamsafar/hotels-booking-data-cleaned-version>

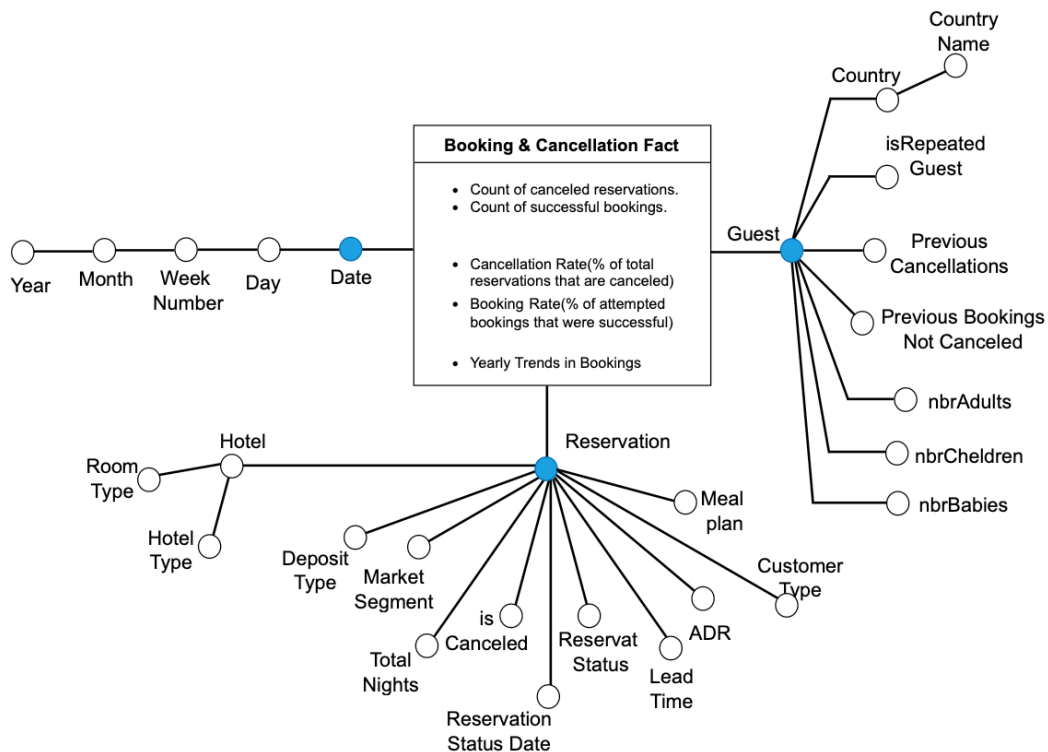
2. **Entity-Relationship (ER) schema :**



3. Attribute tree



4. Dimensional Fact Model (DFM):

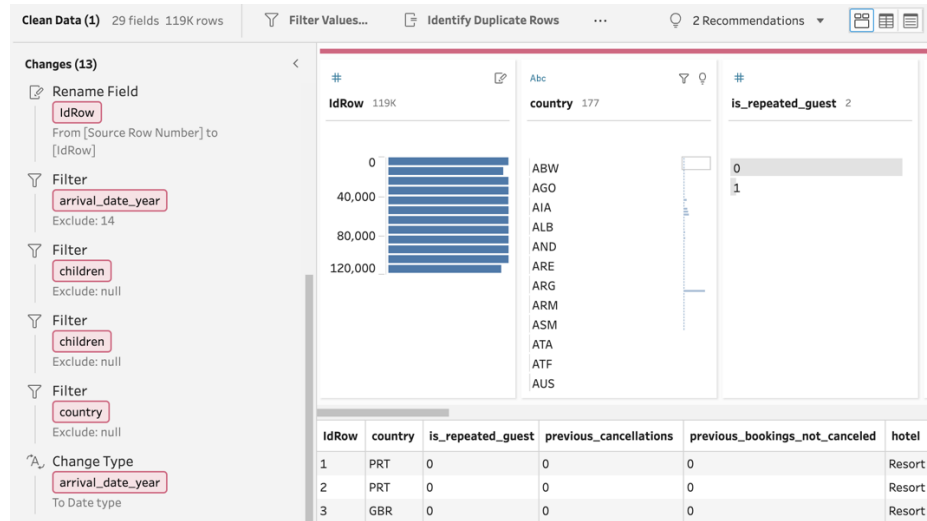


5. My Analysis:

I'm planning to analyze the Cancellation and reservation Rates, and their distribution in terms of deposit type, customer type, market segment, etc. and also i want to see Monthly or yearly trends in cancellations and bookings.

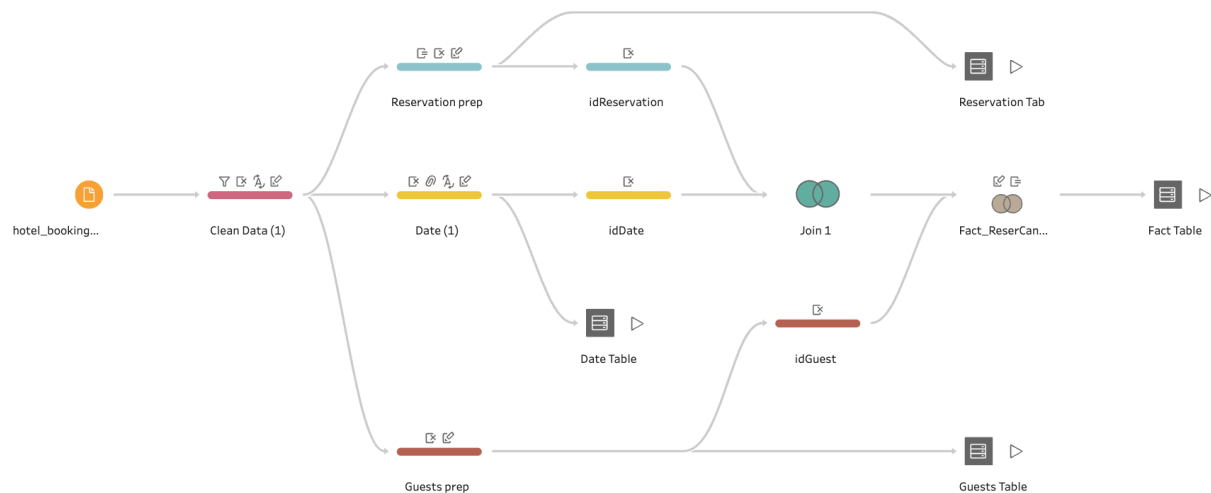
1- Data cleaning and data preparation:

In this step I tried to remove all the unnecessary columns in my data set and keep only those that will help me in my analysis. And also cleaned the NULL Values and change some columns types so that they can match their contents



2- Flow Creation: ETL process:

In step I have designed my ETL process from the reconciled database to the data warehouse. I have used Tableau Prep.



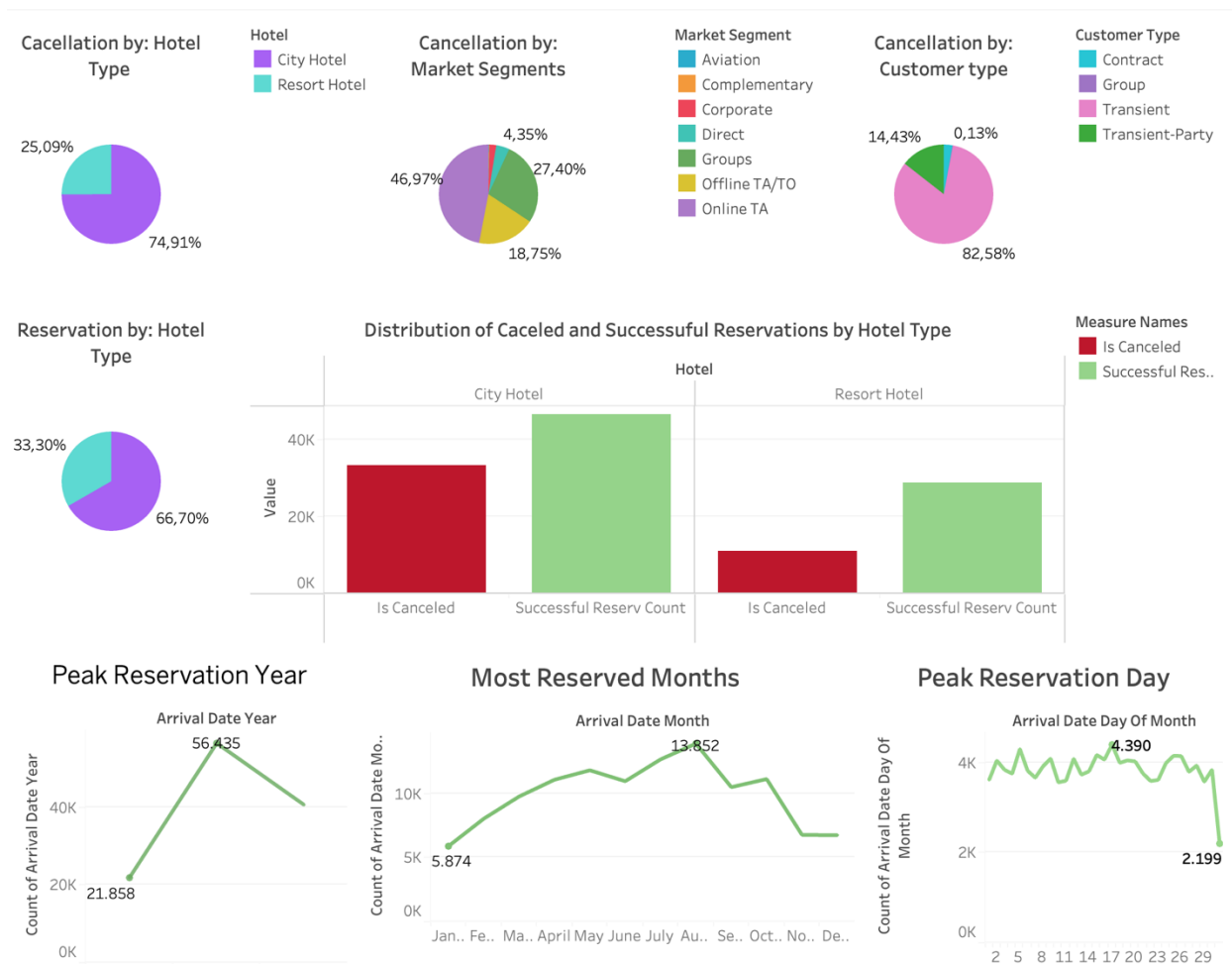
3- Data warehouse:

After creating the flow in Tableau, I have connected tableau to MySQL server so that I can populate my Tables with the needed data.

Hotel_Booking	7	No Deposit	BB	C	Transient	2	0	Check-Out	2015-07-03	0	107	Direct
Tables	8	No Deposit	FB	C	Transient	2	0	Check-Out	2015-07-03	9	103	Direct
Date	9	No Deposit	BB	A	Transient	3	1	Canceled	2015-05-06	85	82	Online TA
Fact_ReserCancel	10	No Deposit	HB	D	Transient	3	1	Canceled	2015-04-22	75	106	Offline TA/TO
Guest	11	No Deposit	BB	E	Transient	4	1	Canceled	2015-06-23	23	123	Online TA
Reservation	12	No Deposit	HB	D	Transient	4	0	Check-Out	2015-07-05	35	145	Online TA
Views	13	No Deposit	BB	D	Transient	4	0	Check-Out	2015-07-05	68	97	Online TA
Stored Procedures	14	No Deposit	HB	G	Transient	4	0	Check-Out	2015-07-05	18	155	Online TA
Functions	15	No Deposit	BB	E	Transient	4	0	Check-Out	2015-07-05	37	95	Online TA
	16	No Deposit	BB	D	Transient	4	0	Check-Out	2015-07-05	68	97	Online TA
	17	No Deposit	BB	E	Contract	4	0	Check-Out	2015-07-05	37	98	Offline TA/TO
	18	No Deposit	BB	A	Transient	1	0	Check-Out	2015-07-02	12	88	Online TA
	19	No Deposit	BB	A	Transient	1	0	Check-Out	2015-07-02	0	107	Corporate
	20	No Deposit	BB	G	Transient	4	0	Check-Out	2015-07-05	7	153	Direct

4- Data Analysis:

Now after creating my data warehouse I moved to data analysis using Tableau Desktop. So first I connected tableau to my dataset and I imported tables, and then I created my Dashboard:

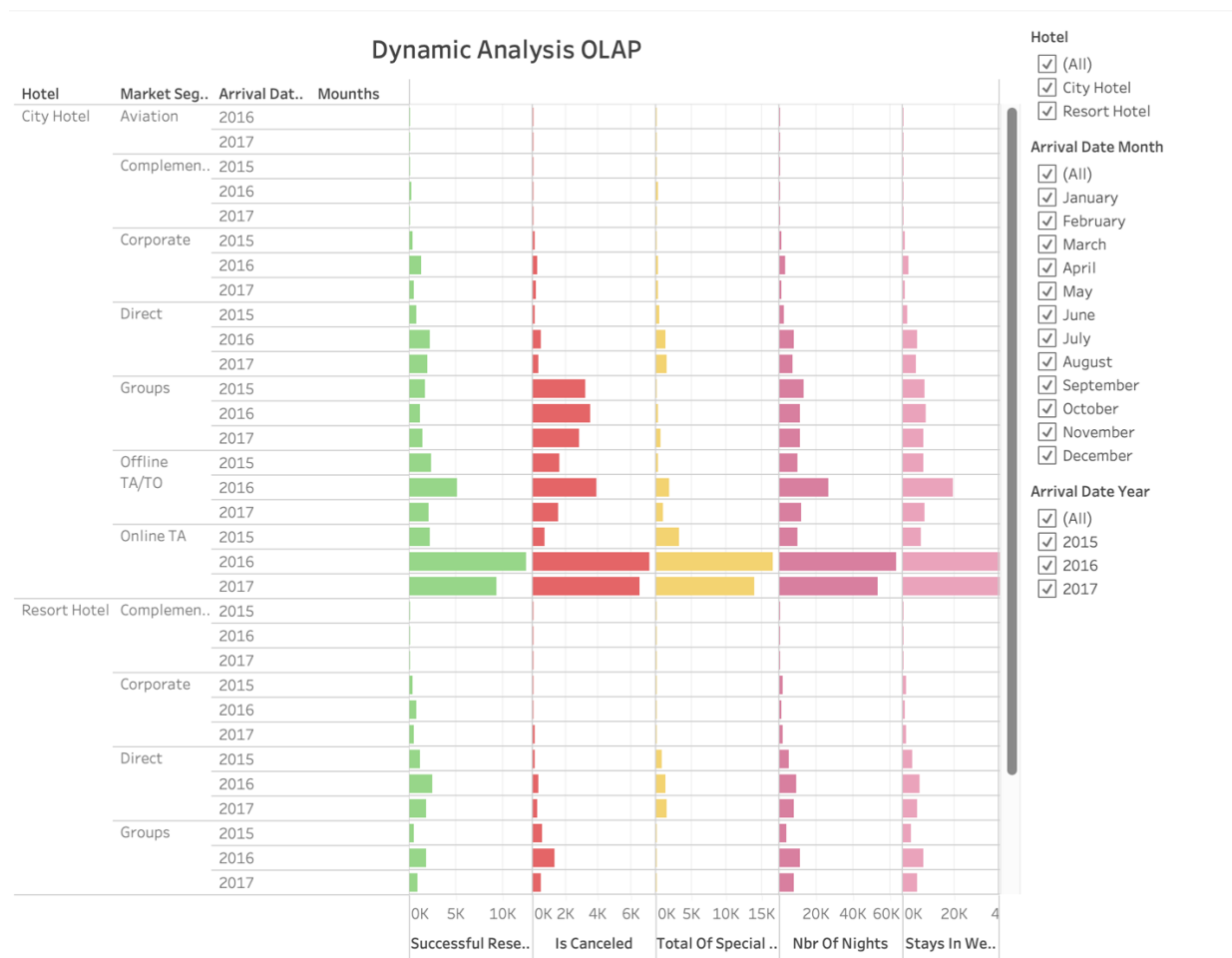


We will try to analyze the cancellation and reservation in resort and city hotel. First of all we can see that the We can see that the cancellation when the resort hotel is less than the cancellation in the city hotel.

And If we go deeper to analyze the reasons of cancelation we can see cancellation based on market segments: we noticed that Travel agencies or groups are the ones that tends to cancel their reservations that is why we should find some solutions to avoid this such as stopping the annulation for the agencies...

And also if we see the customer type we find that the reservations made by transient **individual bookings made by guests who are not part of a group** or do not have any special contractual agreements they tends to cancel more than the reservations that are **made as part of a pre-established agreement or contract between the hotel and a specific entity, such as a company or organization.**

a- And here I have my Dynamic Analysis OLAP:



As a conclusion of this analysis, I can say that the cancellation rate is higher in the city hotel more than the resort hotel and also the reservation that are made by travel agencies or by individual travels that they don't have any contract with the company they tend to cancel their reservations more than the others.

And concerning the peak of reservations it was done in the year 2016 and also for the peak period we can say that the summer period: the months July august September they have the most reservations.