

(Group 2) Data Visualization Mini Project – Hotel Business in Portugal

Objectives

- Understand business performance, a Resort Hotel (H1) in Algarve(法魯) and a City Hotel (H2) in Lisbon (里斯本) covering the period from July 2015 to August 2017
- Diagnose revenue drivers
- Analyze determinants for Average Daily Rate (ADR)
- Investigate the reasons leading to cancelled bookings
- Verify assumptions and hypotheses
- Offer effective insights for revenue maximization and operation optimization

Dataset Description

We used a dataset from Kaggle, in which the dataset is originated from the article “Hotel Booking Demand Datasets” in 2019 and cleaned afterwards. It consists of 32 columns (variables) and 119,390 rows of booking records of a hotel business in Portugal covering from July 2015 to August 2017. Data is stored in mixture of Numeric, Categorical, Text, Date.

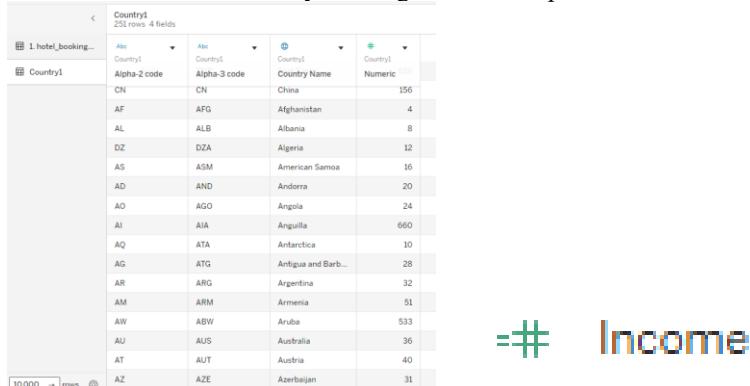
Major information is summarized as below:

Variables	Information
Hotel	Resort Hotel (H1) or City Hotel (H2)
Time Period	July 2015 to August 2017
Customer	country origins, age, background, type of booking, loyalty
Hotel Services Requested	room type, car parking spaces, special requests, booking changes
Booking Channel	Individual / Group Booking e.g. Tour Agent, Tour Operator, Corporate
Payment	Average Daily Rate (ADR), Days of Stay, Deposit Type (No deposit, Non Refund, Refundable)
Final Reservation Status	status and date of recording the status, cancellation

Methodology

Data Preprocessing

We have created new columns for effective analysis. Firstly, country data is stored in abbreviation code, thus lookup was done for data retrieval. Secondly, the revenue formula of each booking was formulated to be “ADR * Days of Stay, and whether booking cancelled at the end”. These were done in Tableau by creating a relationship between tables and a calculated field respectively.



Data Visualization Tool

Tableau was chosen to make charts and build dashboard for effective presentation.

Visualization Approach

Our study basically involved a comprehensive usage of chart types for visualization. On top of that, it is ensured that the advantage of each chart has been taken to deliver core insights effectively. Please refer to appendix for all charts used.

Assumptions

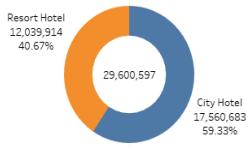
Time Period	Dataset covers from July 2015 to August 2017. The 1st Half of 2015 and 2nd Half 2017 are missing. After reviewing dataset, it is reasonably assumed the hotel is still on a growth stage and no material changes across years.
Room Type	Exactly information other than abbreviation is not provided. After reviewing dataset, it is reasonably assumed a rough classification on size (size grows with alphabetical order).
Occupancy Rate	Not provided. Assume the two hotels don't have full occupancy all the time and there is room to take measures for improvement.
Distribution Channel	Global Distribution System (GDS) : for retail, flexible arrangement & price Channels Other than (GDS) e.g. TA/TO, Direct, Corporate : for special groups, flexible arrangement & price.

Analysis and Findings

There are several directions in our study, which are listed in the following:

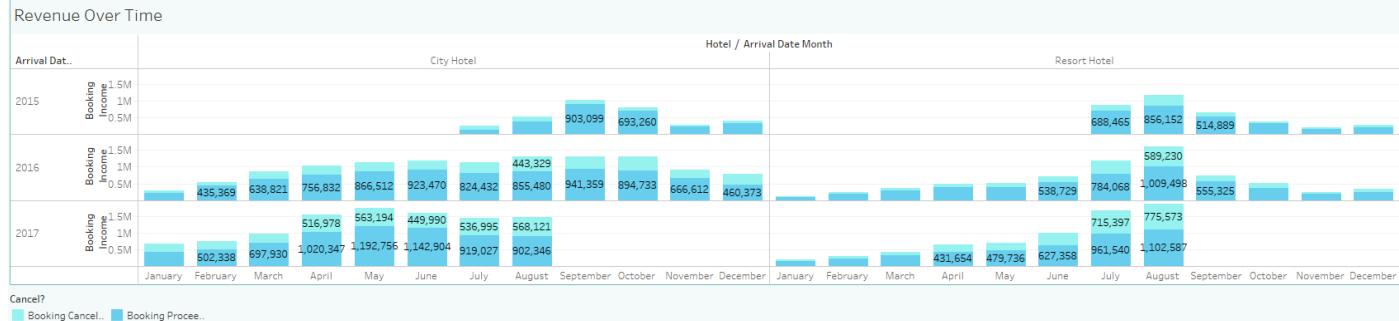
Overview of the hotel business

Revenue by Hotel

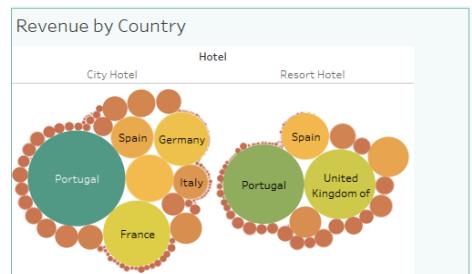


Based on the dataset, city hotel generated total income of \$17,560,683, accounting for 59.33% of the total revenue. In comparison, resort hotel generated total income of \$12,039,914, representing 40.67% of the total revenue.

Revenue Over Time



Hotel income keeps rising. Compared to the previous year, city hotel saw a 29% increase in revenue, while resort hotel experienced a 15% growth. The peak season for city hotel is from April to October, while for resort hotel, it is during July and August.



Revenue by Country

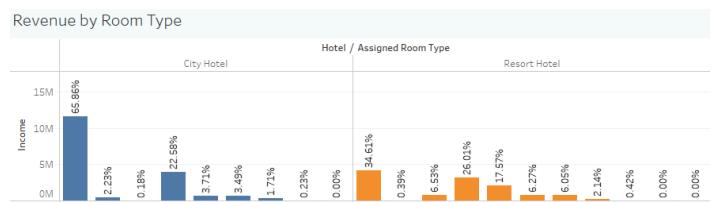
For city hotel, the top three customer origins are as follows: 30.62% from Portugal (local customers), 23.68% from the United Kingdom, and 10.45% from Spain.

For resort hotel, the top three customer origins are: 30.47% from Portugal (local customers), 14.4% from France, and 7.41% from the United Kingdom.

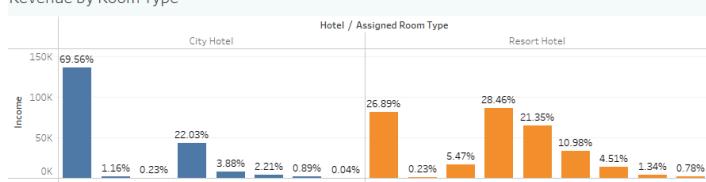
Revenue By Room Type

For city hotel, the majority of customers (65.86%) book room type A, followed by 22.58% who book room type D.

For resort hotel, 35.61% of customers book room type A, 26.01% book room type D, and 17.57% book room type E.



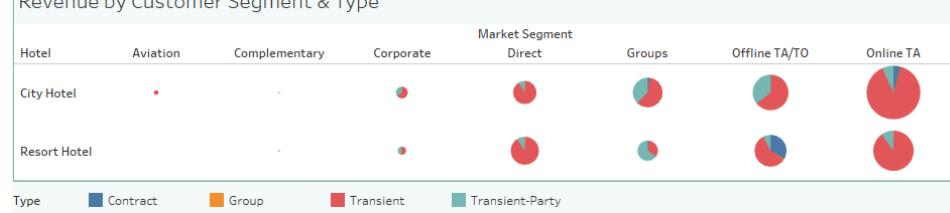
Revenue by Room Type



Revenue By Room Type (Repeat Customer)

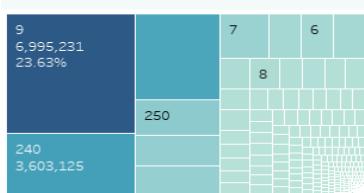
The proportion of repeat customers is very low, accounting for less than 5% of total customers. However, repeat customers are more likely to book better room types during their stays.

Revenue by Customer Segment & Type



The majority of customers are transient. Customers primarily make hotel reservations through online travel agencies (OTAs), offline travel agents/tour operators (TA/TO), and the direct market segment.

Revenue by Agent

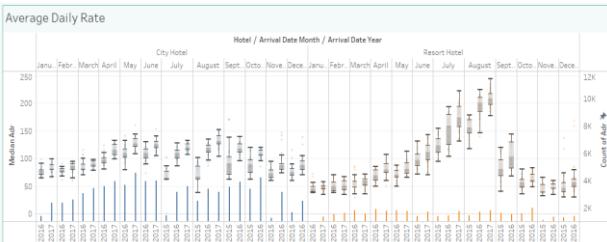


The best agent (#9) generated a total income of \$6,995,231 (23.63%), primarily focusing on city hotels and booking through online travel agencies (OTAs).

The second-best agent (#240) generated a total income of \$3,603,125 (12.17%), primarily focusing on resort hotels and also booking through online travel agencies (OTAs).

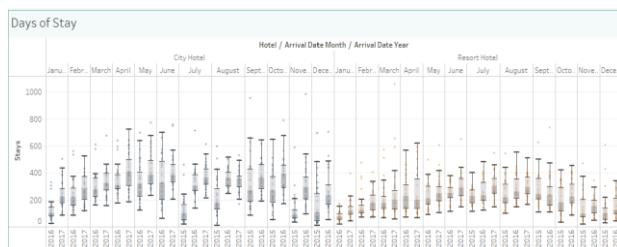
Revenue Drivers

Revenue Drivers are classified into three main areas as below:



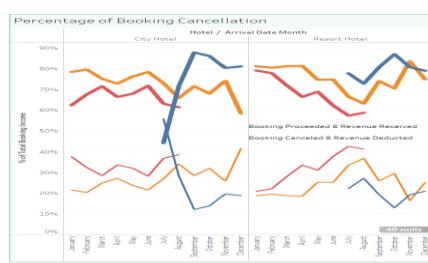
Average Daily Rate (ADR)

ADR and No. of Bookings increase over years and months amid seasonal demand (June, July, August) in the two hotels. It is a good signal which indicates that price and volume both increases.



Days of Stay

The number of days of stay increases over years and months as well. It is a good sign though, but it is also important to note that this factor is limited by length of season and occupancy. Keep it up by constantly improving our service.



Percentage of Booking Cancellation

Although ADR and Days of Stay have witnessed a steady growth over years and months, it is observed that the percentage of revenue not earned due to booking cancellation rose over time in the meantime. This item is definitely worth a deep investigation to sustain a robust growth of the business.

Average Daily Rate (ADR)

3 main hypotheses were made before this study.

- ADR increases with amount of hotel services provided (e.g. room type, catering, other related services)
- ARR is higher for individual booking via online system, whereas it is lower for bulk bookings via direct bargain with hotel
- ADR drops with amount of deposit paid in advance during booking

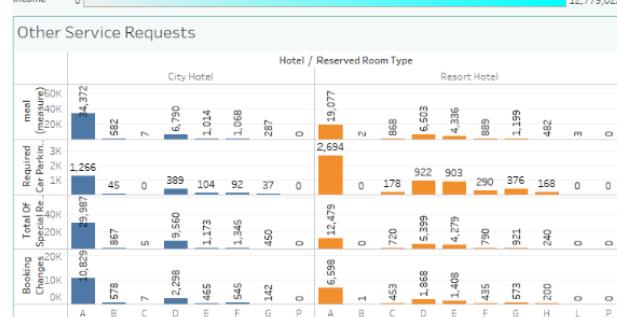


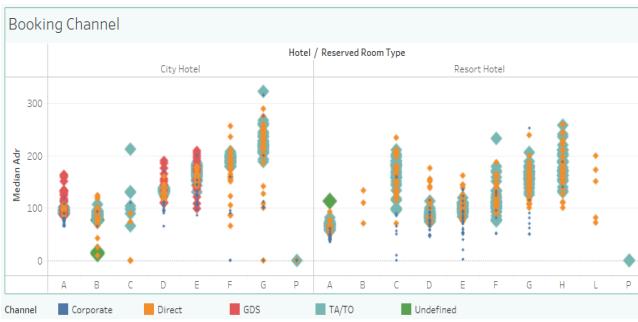
The size of room type is assumed to grows with hotel room alphabetical order. In addition, room type can be further broken down into two segments (A to C) and (D to P) to serve economy class and luxurious class customers respectively.

For segment (D to P), it is observed that ADR increases with hotel room size. On the other hand, segment (A to C) maintains a slight growth among the two hotels.

In addition, it is not suitable to compare the number of bookings and services requested across different room types directly because the number of hotel rooms of their types and customer base are significantly different. However, it is useful to observe the distribution pattern together for the number of bookings and services requested across different room types. Both embody similar distribution which indicates a specific room type is attached with its specific amount of other service requests.

Since ADR increases with hotel room size, it is confident to state ADR increases with amount of hotel services provided.





located at similar level of ADR, it is worth to note that “Corporate” (blue color) is located at relatively lower level of ADR. As corporate should be a more lucrative group from a business perspective, investigation should be taken to question the phenomenon and find ways to see if there is room for improvement.



Booking Cancellation



Cancel By Country

Portugal accounts for 31% of total hotel income but has a significantly higher cancellation rate compared to the overall average. For city hotel, the overall cancellation rate is 42%, while bookings from Portugal show a much higher rate at 65%, which is approximately 55% above average. In resort hotel, the average cancellation rate is 28%, but Portugal's rate reaches 42%, about 50% higher than the hotel average. These elevated cancellation rates from a key income market may affect revenue stability and planning accuracy.

Cancel Rate by Month (Portugal)

For resort hotel, Occupancy rate is inversely proportional to cancellation rate between May to October except July and August. One potential reason for the high cancellation rate is that customers are generally interested in resort hotels and tend to have more flexible schedules, which allows them to book early but cancel or change plans easily as dates approach. To mitigate this, the hotel may consider implementing stricter cancellation policies during the summer season, offering incentives for non-refundable bookings, or introducing flexible rebooking options to retain revenue without discouraging early reservations.

Average Daily Rate of Canceled Each Room Type



Average Daily Rate of Canceled Each Room Type

Room Type A, which is the most popular choice among guests, has the highest cancellation rate. This could mean that although many people are interested in it, the room doesn't stand out enough from other hotels. If customers don't see a clear difference, they might cancel and book somewhere else. It may help to review how Room Type A is promoted and whether its benefits are

In this graph, “GDS” (red color) represents individual booking via online system and others (other colors) represent groups making bulk bookings via direct bargain with hotel.

It is observed that majority of red colors are on top of other colors in *City Hotel - Room Type A, D and E*.

Therefore, it is confident that hypothesis 2 holds within these categories.

While “Direct” (orange color) and “TA/TO” (jade color) are located at similar level of ADR, it is worth to note that “Corporate” (blue color) is located at relatively lower level of ADR. As corporate should be a more lucrative group from a business perspective, investigation should be taken to question the phenomenon and find ways to see if there is room for improvement.

No Deposit (Blue Color), Non Refund (Orange Color), Refundable (Red Color) denote no deposit, full deposit, partial deposit respectively. For segment (D to P), it is observed that majority of other colors are located below blue color which means higher deposit comes with a lower ADR. For segment (A to C), the distribution of other colors spread across blue color in the middle which is due to the fact that these room types have a higher turnover rate where cancelled booking can be replaced by substituted customers rapidly. All in all, hypothesis 3 holds for segment (D to P).

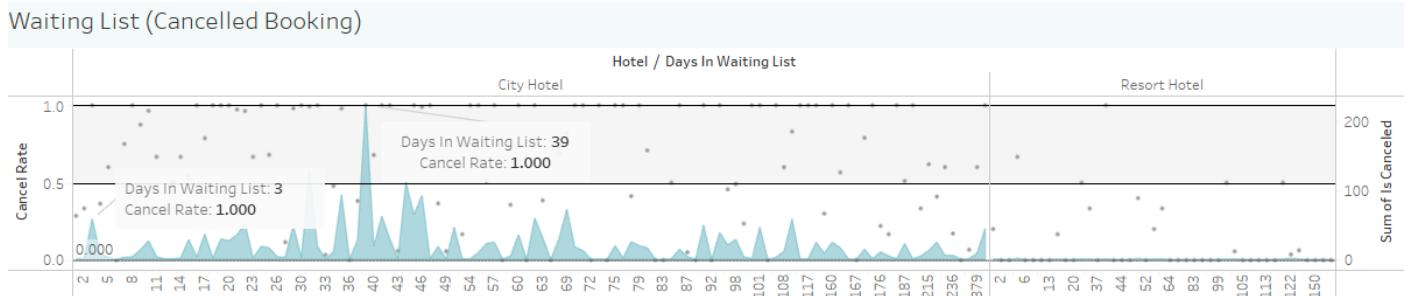
clearly communicated to customers.

We've noticed that as the room type letter goes up (usually meaning better or more expensive rooms), the cancellation rate also increases. This is a trend worth looking into more closely.

There could be a few reasons behind:

- The higher price might make customers think twice or cancel later.
- The room type / price might not attract the right type of customer.
- Customers may not understand the value of these room types before they arrive.

Waiting List (Cancelled Booking)



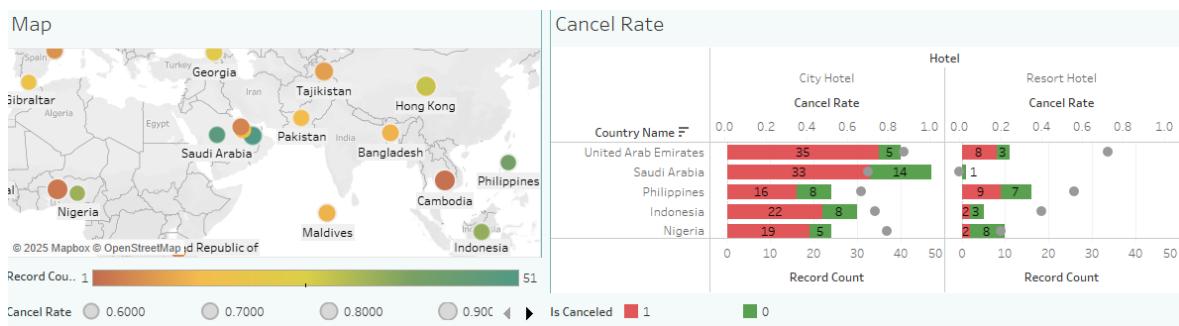
We observed that Day 3 is the first day where the cancellation rate reaches 100%, and Day 39 is the day with the highest number of cancellations overall. These two dates may represent key psychological decision points for customers who are still uncertain about their bookings. To reduce cancellations, it is recommended to secure reservation confirmations before these dates, through follow-up emails, reminders, or special offers that encourage commitment.

Deposit Type



Non-refundable booking policies can help reduce cancellations by creating a sense of commitment. When customers know they can't get their money back, they're less likely to cancel their reservation. The idea of losing the payment acts as a kind of mental "barrier," making customers more likely to follow through with their plans. By using non-refundable options, hotels can encourage guests to keep their bookings. Of course, it's important to strike a balance to make sure it's fair and doesn't turn customers away in the first place.

Cancel Rate by Country (Cancel Rate >=0.6)



We've noticed that customers from wealthy regions often book our hotel, but their cancellation rates are unusually high. This indicates that while we're able to reach a diverse range of customers globally, we may not have enough strategies in place to retain them. It's important to investigate whether the information these customers receive about our hotel is clear and sufficient.

Challenges and Limitations

Challenges	<ul style="list-style-type: none"> Hotel operation varies across scale, brand and location of the hotel. Those factors have to be seriously considered while doing comparison and generating reasonable insights.
Limitations of the approach or dataset	<ul style="list-style-type: none"> The value of hotel property might dictate the performance of the hotel e.g. whether the hotel is situated in a superior location Detailed information is not provided e.g. booking information of 1st Half 2015 and 2nd Half 2017, customer gender, room type, occupancy rate, cancellation handling procedure Data source timeframe is pre-pandemic, conditions might have changed after that if the analysis outcome is to apply to a post-pandemic business.
Potential areas for future improvement	<ul style="list-style-type: none"> In-person inspection e.g. Questionnaires about further improvement on hotel facilities and service. Visit the hotels and directly liaise with management to understand the business Take into account external factors given sufficient information about the hotel is provided e.g. competitors, local regulations, local economy

Future Work

- Description of any additional ideas or approaches that were not implemented
 - Deep investigation of customer segments, hotel room types
 - Take into account more operations and related services of hotel business for profound analysis
e.g. discount to travel agencies, loyalty program to customers, early bird discounts to booking, cancellation fee
- Reasons for not implementing these ideas
 - Information provision: not all information is fairly disclosed to the public and it is also hard to collect
 - Time constraint: deep investigation on hotel facilities, services is time-consuming
 - Complexity: hotel operation varies across scale and brand of the hotel, which is hard for direct comparison

Conclusion

Summary of the project's objectives and achievements	<p><u>Objectives & Achievements</u></p> <ul style="list-style-type: none"> Understood business performance, a Resort Hotel (H1) in Algarve(法魯) and a City Hotel (H2) in Lisbon (里斯本), covering the period from July 2015 to August 2017 Diagnosed revenue drivers Analyzed determinants for Average Daily Rate (ADR) Investigated reasons leading to cancelled bookings Verified assumptions and hypotheses Offered effective insights for revenue maximization and operation optimization
Key takeaways and insights	<p><u>Overview of the hotel business</u></p> <ul style="list-style-type: none"> City hotel vs resort hotel Time-series Customer (country origin, age & loyalty, segment & type) Room type & related services Agent performance <p><u>Revenue Drivers</u></p> <ul style="list-style-type: none"> ADR Days of stay Number of bookings Percentage of booking cancellation <p><u>Average Daily Rate (ADR)</u></p> <ul style="list-style-type: none"> Room type & other service requests Booking channel Deposit type <p><u>Booking Cancellation</u></p> <ul style="list-style-type: none"> Country Month Room type Deposit type Waiting List Assignment of Room Type

Work Distribution

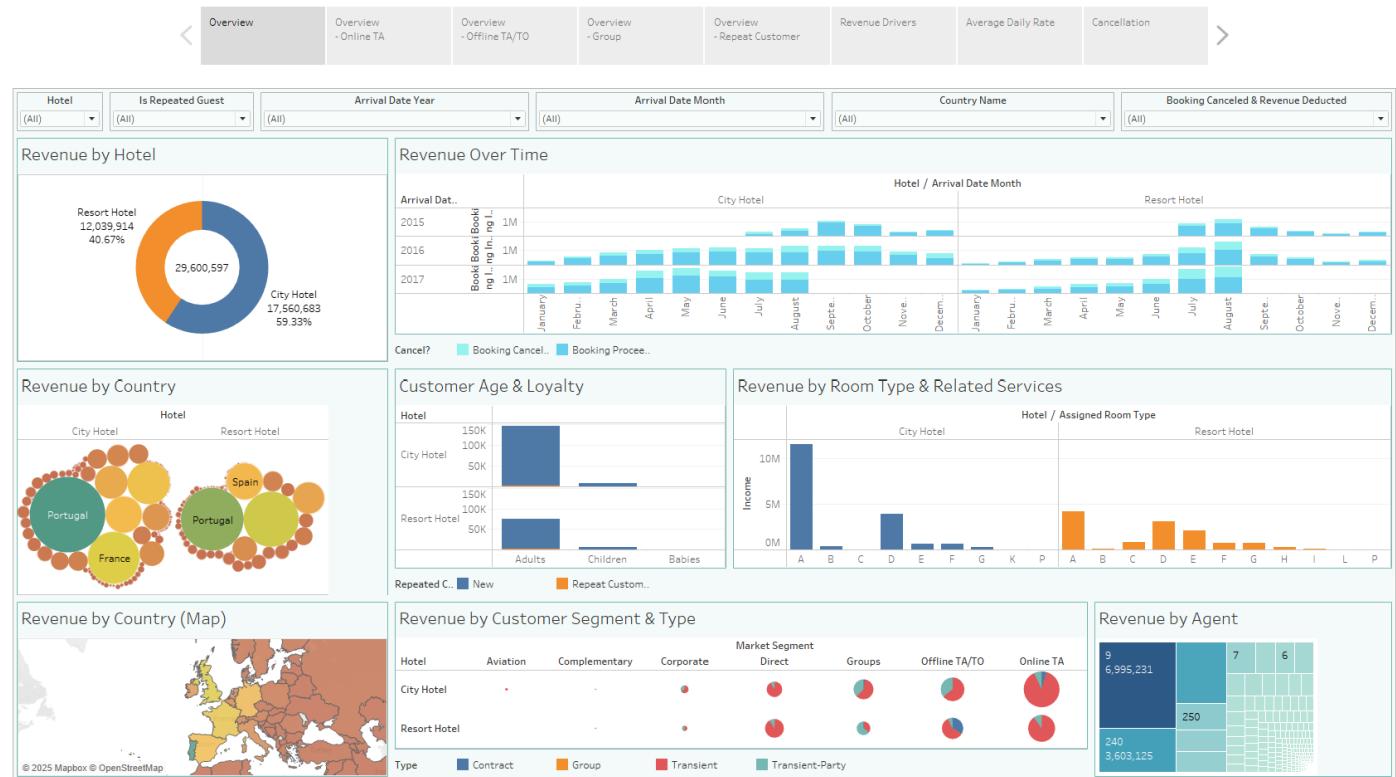
Even distribution among team members

Appendix

Dashboard in Full Version

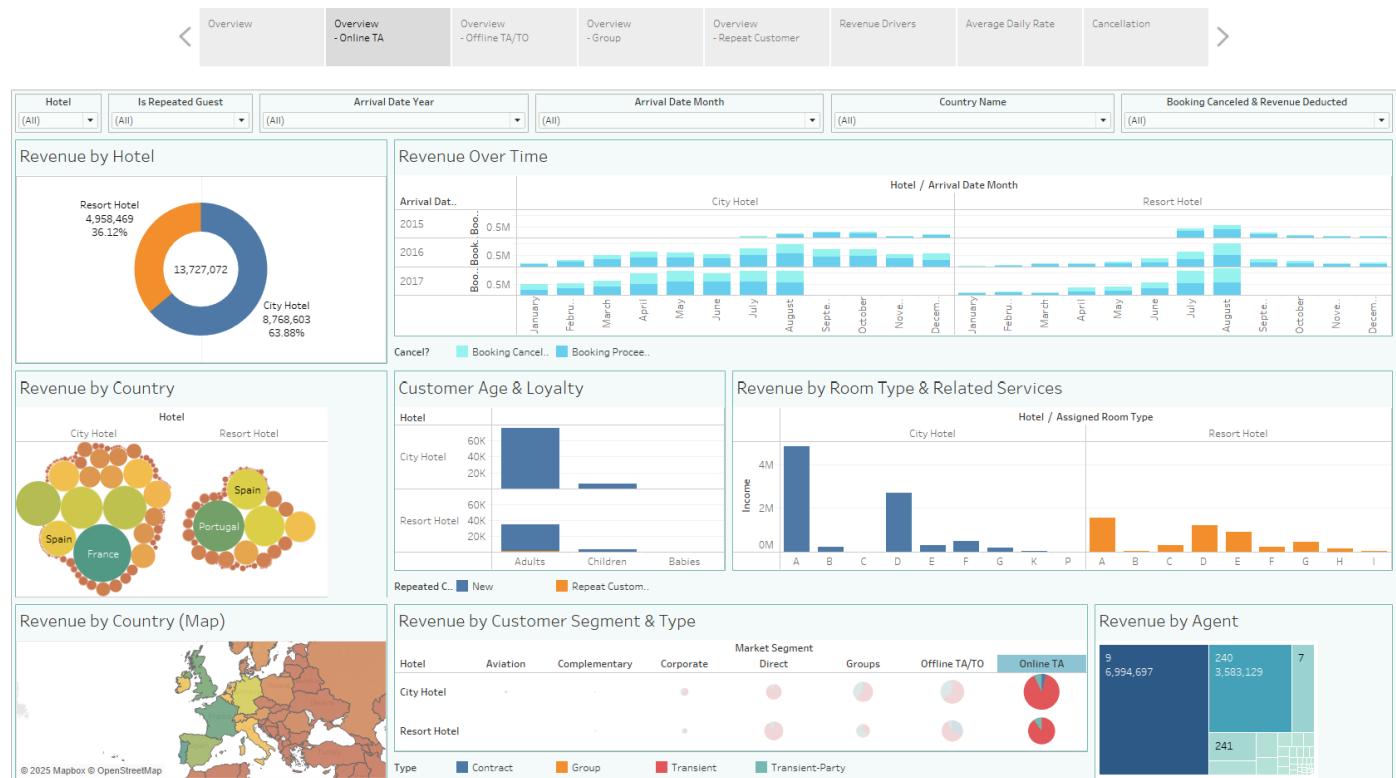
Tab 1 - Overview

Portugal Hotel Business



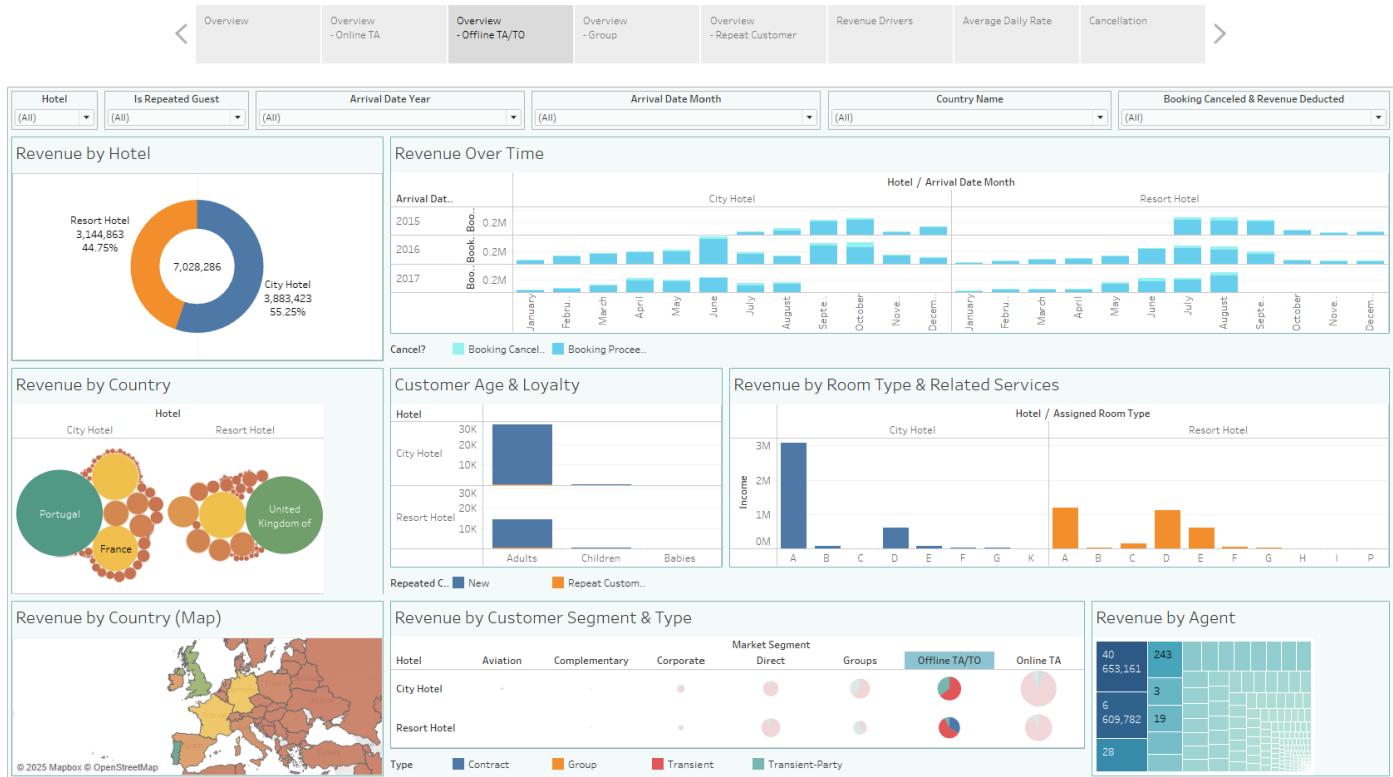
Tab 2 – Overview (Online TA)

Portugal Hotel Business



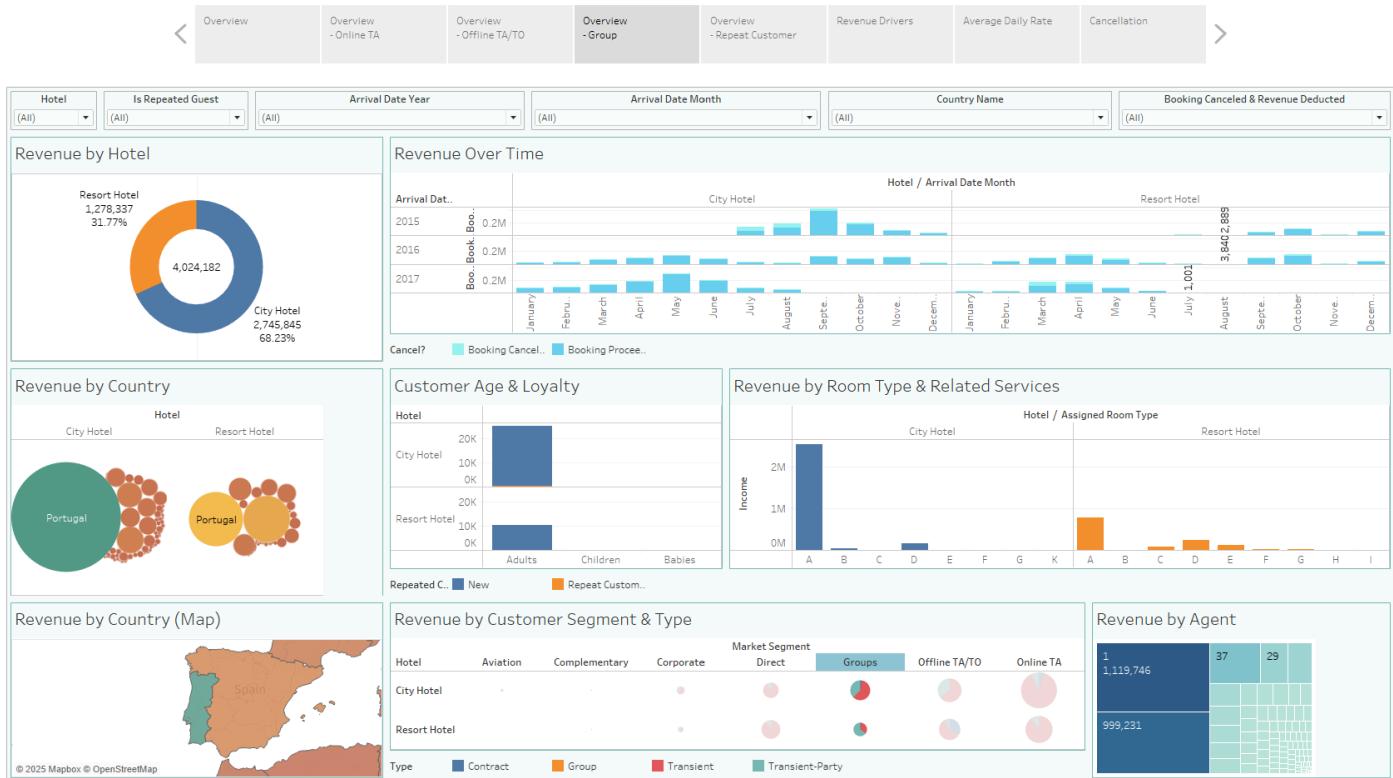
Tab 3 – Overview (Offline TA/TO)

Portugal Hotel Business



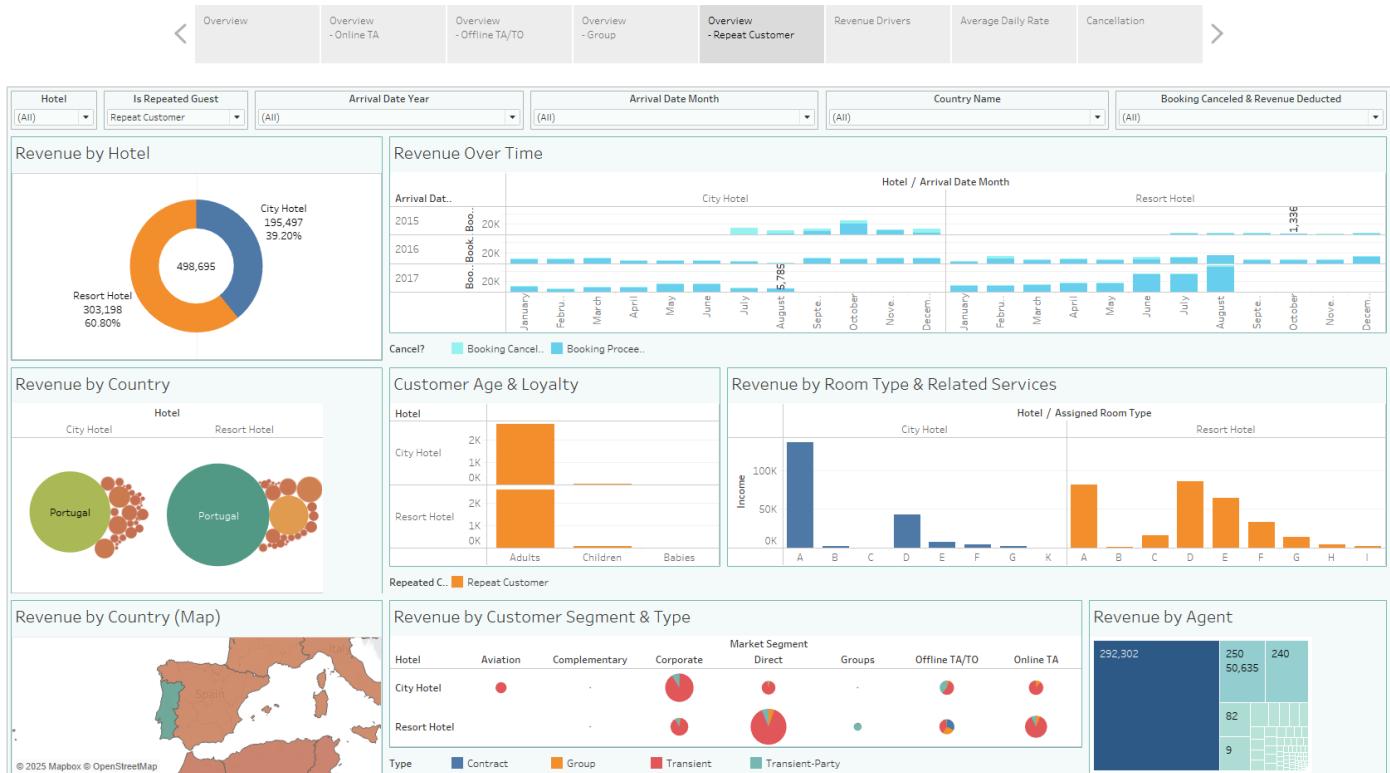
Tab 4 – Overview (Group)

Portugal Hotel Business



Tab 5 – Overview (Repeat Customer)

Portugal Hotel Business



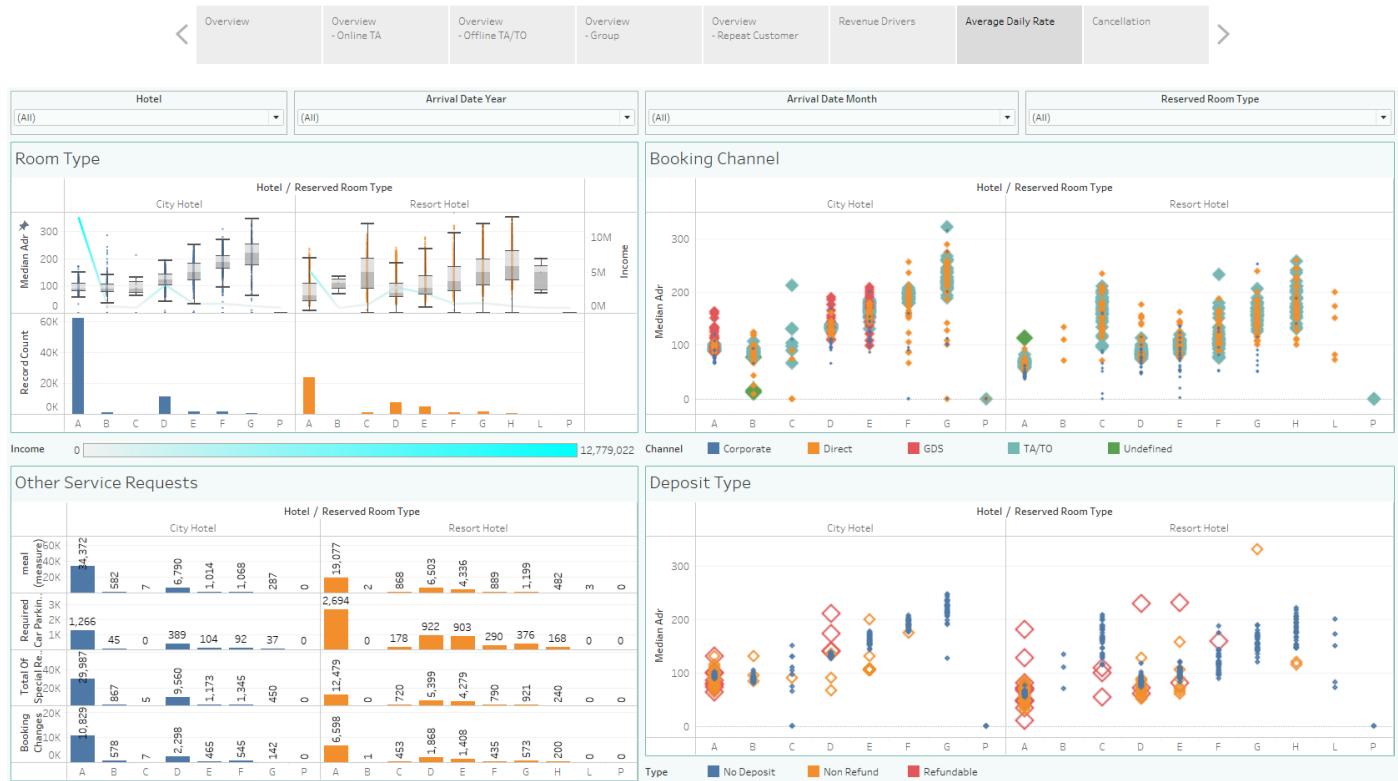
Tab 6 – Revenue Drivers

Portugal Hotel Business



Tab 7 – Average Daily Rate

Portugal Hotel Business



Tab 8 – Cancellation

Portugal Hotel Business

