Hotel Reservation Analysis

With PostgreSQL

By Rashidat Ambali

Project Overview

The hotel industry relies on data to make informed decisions and provide a better guest experience. In this project, I analyzed hotel reservation dataset to gain insights into guest preferences, booking trends, and other key factors that impact the hotel's operations.

Data Description

The dataset consists of 12 columns (attributes), and 700 rows (observations). The column names include the following:

- Booking_ID: A unique identifier for each hotel reservation.
- no_of_adults: The number of adults in the reservation.
- no_of_children: The number of children in the reservation.
- no_of_weekend_nights: The number of nights in the reservation that fall on weekends.
- no_of_week_nights: The number of nights in the reservation that fall on weekdays.
- type_of_meal_plan: The meal plan chosen by the guests.
- room_type_reserved: The type of room reserved by the guests.
- lead_time: The number of days between booking and arrival.
- arrival_date: The date of arrival.
- market_segment_type: The market segment to which the reservation belongs.
- avg_price_per_room: The average price per room in the reservation.
- booking_status: The status of the booking.

Data Preparation

- **Data Collection:** I collected the dataset file via the link provided by Mentorness, opened it in Excel, and conducted a preliminary review to understand the data structure, data types, and content.

- **Data Cleaning:** I converted the data types of some columns to their appropriate data types using Excel, and saved the Excel file as a csv file. No missing values, no duplicates, and no no inconsistencies in the dataset.

Methodology

The following steps outline the methodology employed in this project.

- **Data Cleaning:** A clean dataset was provided. However, the data type of some columns were changed from ‘General’ to their appropriate data type using Microsoft Excel, then saved as a CSV file.

- **Data Analysis:** After ensuring data integrity, I created a database called “Mentorness” on PostgreSQL, created a table named “Hotel Reservation”, imported the CSV file into the table, and queried and analyze the data, by answering the following questions about the dataset.

Methodology Cont’d

1. What is the total number of reservations in the dataset?
2. Which meal plan is the most popular among guests?
3. What is the average price per room for reservations involving children?
4. How many reservations were made for the year 20XX (replace XX with the desired year)?
5. What is the most commonly booked room type?
6. How many reservations fall on a weekend (no_of_weekend_nights > 0)?
7. What is the highest and lowest lead time for reservations?
8. What is the most common market segment type for reservations?
9. How many reservations have a booking status of "Confirmed"?
10. What is the total number of adults and children across all reservations?
11. What is the average number of weekend nights for reservations involving children?
12. How many reservations were made in each month of the year?
13. What is the average number of nights (both weekend and weekday) spent by guests for each room type?
14. For reservations involving children, what is the most common room type, and what is the average price for that room type?
15. Find the market segment type that generates the highest average price per room.

Data Analysis

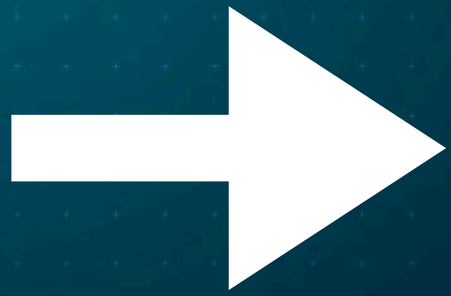

Table Creation

```
CREATE TABLE Hotel_Reservation (
    Booking_ID varchar(100),
    no_of_adults int,
    no_of_children int,
    no_of_weekend_nights int,
    no_of_week_nights int,
    type_of_meal_plan varchar(100),
    room_type_reserved varchar(100),
    lead_time int,
    arrival_date date,
    market_segment_type text,
    avg_price_per_room numeric,
    booking_status varchar(100)
);
```

Table Content View

```
SELECT *
FROM Hotel_Reservation
```

Output | Messages | Notifications

booking_id character varying (100)	no_of_adults integer	no_of_children integer	no_of_weekend_nights integer	no_of_week_nights integer	type_of_meal_plan character varying (100)	room_type_reserved character varying (100)
INN00001	2	0	1	2	Meal Plan 1	Room_Type 1
INN00002	2	0	2	3	Not Selected	Room_Type 1
INN00003	1	0	2	1	Meal Plan 1	Room_Type 1
INN00004	2	0	0	2	Meal Plan 1	Room_Type 1
INN00005	2	0	1	1	Not Selected	Room_Type 1

al rows: 700 of 700 | Query complete 00:00:00.221 | Ln 16, Col 1

What is the total number of reservations in the dataset?

Total Number of Reservations
700

```
21
22  ---1. What is the total number of reservations in the dataset?
23  SELECT COUNT(*) AS Total_Reservations
24  FROM Hotel_Reservation
25
26
```

Data Output | Messages | Notifications

	total_reservations bigint
1	700

Which meal plan is the most popular among guests?

Most Popular Meal Plan
(Meal Plan 1)

```
28  ---2. Which meal plan is the most popular among guests?
29  SELECT type_of_meal_plan AS meal_plan, COUNT(type_of_meal_plan)
30  AS meal_plan_count
31  FROM Hotel_Reservation
32  GROUP BY meal_plan
33  ORDER BY  meal_plan_count DESC
34  LIMIT 1
35
36
```

Data Output | Messages | Notifications

	meal_plan character varying (100)	meal_plan_count bigint
1	Meal Plan 1	527

What is the average price per room for reservations involving children?

Average Price per Room (Children)

```sql
---3. What is the average price per room for reservations involving children?
SELECT Booking_ID, no_of_children, avg_price_per_room
FROM Hotel_Reservation
WHERE no_of_children > 0;
```

Data Output | Messages | Notifications

	booking_id character varying (100)	no_of_children integer	avg_price_per_room numeric
1	INN00033	2	82.44
2	INN00061	2	258.00
3	INN00081	2	159.30
4	INN00096	1	130.50
5	INN00100	2	156.90
6	INN00115	2	184.24

How many reservations where made in the year 20XX (replace XX with the desired year)?

Total Reservations in 2018

577

```
42
43  ---4. How many reservations were made for the year 20XX (replace XX with the desired year)?
44  SELECT EXTRACT(YEAR FROM arrival_date) AS year, COUNT(*) AS total_reservations
45  FROM Hotel_Reservation
46  WHERE EXTRACT(YEAR FROM arrival_date) = 2018
47  GROUP BY year;
48
49
50
51
52
```

Data Output | Messages | Notifications

	year numeric	total_reservations bigint
1	2018	577

What is the most commonly booked room type?

Most Commonly Booked Room Type (Room_Type 1)

```sql
---5. What is the most commonly booked room type?
SELECT room_type_reserved, COUNT(*) AS total_reservations
FROM Hotel_Reservation
GROUP BY room_type_reserved
ORDER BY total_reservations DESC
LIMIT 1;
```

Data Output | Messages | Notifications

	room_type_reserved character varying (100)	total_reservations bigint
1	Room_Type 1	534

How many reservations fall on a weekend (no_of_weekend_nights > 0)?

Weekend Reservations
383

```
65  ---6. How many reservations fall on a weekend (no_of_weekend_nights > 0)?
66  SELECT COUNT(*) AS weekend_reservations
67  FROM Hotel_Reservation
68  WHERE no_of_weekend_nights > 0;
69
70
71
72
73
74
```

Data Output | Messages | Notifications

	weekend_reservations bigint
1	383

What is the highest and lowest lead time for reservations?

Highest Lead Time
443
Lowest Lead Time
0

```
74
75  ---7. What is the highest and lowest lead time for reservations?
76  SELECT MAX(lead_time) AS highest_lead_time, MIN(lead_time) AS lowest_lead_time
77  FROM Hotel_Reservation
78
79
```

Data Output | Messages | Notifications

	highest_lead_time integer	lowest_lead_time integer
1	443	0

What is the most common market segment type for reservations?

Most Common Market Segment (Online)

```sql
---8. What is the most common market segment type for reservations?
SELECT market_segment_type, COUNT(market_segment_type) AS market_segment_count
FROM Hotel_Reservation
GROUP BY market_segment_type
ORDER BY  market_segment_count DESC
LIMIT 1;
```

Data Output | Messages | Notifications

	market_segment_type text	market_segment_count bigint
1	Online	518

How many reservations have a booking status of “Confirmed”?

Confirmed Reservations
493

```
95
96  ---9. How many reservations have a booking status of "Confirmed"?
97  SELECT booking_status, COUNT(*) AS booking_status_count
98  FROM Hotel_Reservation
99  GROUP BY booking_status
100 HAVING booking_status = 'Not_Canceled';
101
102
103
```

Data Output | Messages | Notifications

	booking_status character varying (100)	booking_status_count bigint
1	Not_Canceled	493

What is the total number of adults and children across all reservations?

```sql
---10. What is the total number of adults and children across all reservations?
SELECT SUM(no_of_adults) AS total_adults, SUM(no_of_children) AS total_children
FROM Hotel_Reservation;
```

Data Output | Messages | Notifications

	total_adults bigint	total_children bigint
1	1316	69

Total Adults
1316
Total Children
69

What is the average number of weekend nights for reservations involving children?

Average Weekend Reservations (Children)

1

```sql
---11. What is the average number of weekend nights for reservations involving children?
SELECT ROUND(AVG(no_of_weekend_nights), 2) AS average_weekend_reservations
FROM Hotel_Reservation
WHERE no_of_children > 0;
```

Data Output | Messages | Notifications

	average_weekend_reservations numeric
1	1.00

How many reservations were made in each month of the year?

```sql
---12. How many reservations were made in each month of the year?
SELECT EXTRACT(MONTH FROM arrival_date) AS Reservation_Month,
    COUNT(*) Reservation_Count
FROM Hotel_Reservation
GROUP BY Reservation_Month
ORDER BY Reservation_Month;
```

Data Output | Messages | Notifications

	reservation_month numeric	reservation_count bigint
1	1	11
2	2	28
3	3	52
4	4	67
5	5	55
6	6	84

Highest Reservation (October)

Lowest Reservation (January)

What is the average number of nights (both weekend and weekday) spent by guests for each room type?

Average Weekday and Weekend nights Per Room Type

```
135  ---13. What is the average number of nights (both weekend and weekday) spent by guests for each room type?
136  SELECT room_type_reserved, ROUND(AVG(no_of_week_nights), 4) AS Average_weekday,
137      ROUND(AVG(no_of_weekend_nights), 4) AS Average_weekend
138  FROM Hotel_Reservation
139  GROUP BY room_type_reserved
140  ORDER BY room_type_reserved;
141
142
```

Data Output | Messages | Notifications

	room_type_reserved character varying (100)	average_weekday numeric	average_weekend numeric
1	Room_Type 1	2.0899	0.7884
2	Room_Type 2	2.0000	1.0000
3	Room_Type 4	2.7077	1.0923
4	Room_Type 5	2.5000	0.0000
5	Room_Type 6	2.5556	1.0556
6	Room_Type 7	1.6667	1.0000

For reservations involving children, what is the most common room type, and what is the average price for that room type?

Most Common Room Type (Children)

(Room_Type 1)

Average Price Per Room Type Count

123.1229

```
142
143 /*14. For reservations involving children, what is the most common room type, and what is the
144 average price for that room type?*/
145
146 SELECT room_type_reserved, Count(room_type_reserved) AS room_type_count,
147     AVG(avg_price_per_room) AS Avg_price_per_room_type_count
148 FROM Hotel_Reservation
149 WHERE no_of_children > 0
150 GROUP BY room_type_reserved
151 ORDER BY room_type_count DESC
152 LIMIT 1;
```

Data Output | Messages | Notifications

	room_type_reserved character varying (100)	room_type_count bigint	avg_price_per_room_type_count numeric
1	Room_Type 1	24	123.1229166666666667

Find the market segment type that generates the highest average price per room.

Market Segment with the Highest Average Price per Room (Online)

```sql
---15. Find the market segment type that generates the highest average price per room.
SELECT market_segment_type, AVG(avg_price_per_room) AS Avg_price_per_market_segment
FROM Hotel_Reservation
GROUP BY market_segment_type
ORDER BY Avg_price_per_market_segment DESC
LIMIT 1;
```

Data Output | Messages | Notifications

	market_segment_type text	avg_price_per_market_segment numeric
1	Online	112.4552123552123552

Insights

- The most popular meal plan among guests is "meal plan 1".
- There were 577 reservations made in the year 2018. This indicates that a significant portion of the reservations (82.57% of the total reservations) were made in 2018.
- The most commonly booked room type is "room-type 1".
- There were 383 reservations that included weekend nights.
- The highest lead time is 443 days, and the lowest lead time is 0 days. The high lead time indicates some guests plan their stays well in advance, while the zero lead time shows last-minute bookings.
- The most common market segment is "online".
- There are 493 confirmed reservations.
- The total number of adults is 1316, and the total number of children is 69.
- The average number of weekend nights for reservations involving children is 1.
- October has the highest number of reservations (103), while January has the lowest (11).
- The market segment with the highest average price per room is "online".

Recommendations

The hotel could:

- Explore why “meal_plan 1” is popular and consider marketing it more or replicating its features in other plans.
- Investigate what contributed to the higher number of reservations in 2018, such as marketing campaigns, economic factors, or seasonal trends.
- Increase availability or features of "room-type 1" to attract more bookings.
- Develop special weekend packages or promotions to attract more guests during weekends.
- Tailor marketing strategies for both early planners and last-minute bookers.
- Enhance online booking platforms and digital marketing efforts.
- Analyze and address factors leading to cancellations to increase confirmed bookings.
- Consider offering more family-oriented amenities and promotions to attract more families.
- Boost marketing efforts during low reservation months like January, and capitalize on high-demand months like October.
- Focus on exclusive and premium packages for online bookings to maximize revenue.

Thank You!

Do you have any question/suggestion?

ambalirasheedat@gmail.com
www.linkedin.com/in/rashidat-ambali-698726101


slidesgo