Hotel Reservation Analysis

Presented By Deepa Saini

Project Overview

The hotel industry relies on data to make informed decisions and enhance the overall guest experience. In this project, we delve into a comprehensive hotel reservation dataset to extract meaningful insights. Our goal is to uncover guest preferences, Identify booking trends and discover key factors influencing how the hotel operates.

Objective

Our objective is to leverage SQL for in-depth exploration and analysis of the dataset. By addressing specific queries related to the dataset , our goal is to reveal patterns that will guide strategic decisions and optimize the overall performance of the hotel.

Tools Used

Dataset Overview

The dataset includes the following columns:

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.

Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	room_type_reserved	lead_time	arrival_date	market_segment_type	avg_price_per_room	booking_status
INN00001	2	0	1	2	Meal Plan 1	Room_Type 1	224	2017-10-02	Offline	65	Not_Canceled
INN00002	2	0	2	3	Not Selected	Room_Type 1	5	2018-11-06	Online	106.68	Not_Canceled
INN00003	1	0	2	1	Meal Plan 1	Room_Type 1	1	2018-02-28	Online	60	Canceled
INN00004	2	0	0	2	Meal Plan 1	Room_Type 1	211	2018-05-20	Online	100	Canceled
INN00005	2	0	1	1	Not Selected	Room_Type 1	48	2018-04-11	Online	94.5	Canceled
INN00006	2	0	0	2	Meal Plan 2	Room_Type 1	346	2018-09-13	Online	115	Canceled

Data Exploration with SQL

```sql
-- Viewing the Dataset
USE projects;
SELECT * FROM hotel_data;
```

```sql
-- Size of the Dataset
SELECT COUNT(*) AS total_row FROM hotel_data;
```

total_row
700

```sql
-- Exploring Date Range
SELECT MIN(arrival_date) AS start_date, MAX(arrival_date) AS end_date FROM hotel_data;
```

start_date	end_date
2017-07-01	2018-12-31

```sql
DESCRIBE hotel_data;
-- (Identified issue: arrival_date data type is incorrect)
```

```sql
-- Correcting Data Type for arrival_date
SET SQL_SAFE_UPDATES = 0;
UPDATE hotel_data
SET arrival_date = STR_TO_DATE(arrival_date, '%d-%m-%Y');
ALTER TABLE hotel_data
MODIFY arrival_date DATE ;
```

Field	Type	Null	Key	Default
Booking_ID	text	YES		NULL
no_of_adults	int	YES		NULL
no_of_children	int	YES		NULL
no_of_weekend_nights	int	YES		NULL
no_of_week_nights	int	YES		NULL
type_of_meal_plan	text	YES		NULL
room_type_reserved	text	YES		NULL
lead_time	int	YES		NULL
arrival_date	date	YES		NULL
market_segment_type	text	YES		NULL
avg_price_per_room	dou...	YES		NULL
booking_status	text	YES		NULL

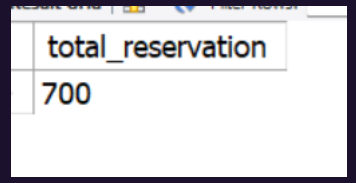
Insights:
The dataset consists of 700 rows and 12 columns, spanning from 2017 to 2018.

Data Analysis Queries

#Query1 :

What is the total number of reservations in the dataset?

```
SELECT COUNT(Booking_ID) AS total_reservation
FROM hotel_data;
```



total_reservation
700

Insights:
There are 700 reservations in the dataset.

Data Analysis Queries

#Query2 :

Which meal plan is the most popular among guests?

```sql
SELECT type_of_meal_plan, COUNT(type_of_meal_plan) AS total_count FROM hotel_data
GROUP BY type_of_meal_plan
ORDER BY COUNT(type_of_meal_plan) DESC
LIMIT 1;
```

type_of_meal_plan	total_count
Meal Plan 1	527

Insights:
Meal Plan 1 stands out as the most popular choice among guests.

Data Analysis Queries

#Query3:

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

```
SELECT ROUND(AVG(avg_price_per_room),1) AS avg_price_per_room_for_children
FROM hotel_data
WHERE no_of_children >0 ;
```

avg_price_per_room_for_children
144.6

Insights:
Reservations involving children have an average room price of 144.6.

Data Analysis Queries

#Query4:

How many reservations were made in each year?

```
SELECT YEAR(arrival_date) AS year  ,COUNT(*) AS total_reservations FROM hotel_data
GROUP BY YEAR(arrival_date)
ORDER BY YEAR(arrival_date) ASC ;
```

year	total_reservations
2017	123
2018	577

Insights:
In 2017, there were 123 reservations. Subsequently, in 2018, the reservations experienced a significant increase, indicating positive growth.

Data Analysis Queries

#Query5:

What is the most commonly booked room type?

```
SELECT room_type_reserved, COUNT(room_type_reserved) AS total_booking FROM hotel_data
GROUP BY room_type_reserved
ORDER BY COUNT(room_type_reserved) DESC
LIMIT 1;
```

room_type_reserved	total_booking
Room_Type 1	534

Insights:
Room Type 1 is the guests' top choice for bookings.

Data Analysis Queries

#Query6:

How do guest reservations vary between weekend and weekday nights?

```sql
SELECT COUNT(*) AS total_weekend_reservations
FROM hotel_data
WHERE no_of_weekend_nights > 0 ;
```

total_weekend_reservations
383

```sql
-- weekday reservation
SELECT COUNT(*) AS total_weekday_reservations
FROM hotel_data
WHERE no_of_week_nights > 0 ;
```

total_weekday_reservations
656

Insights:

The higher number of reservations for weekday nights (656) compared to weekend nights (383) suggests a guest preference for staying during weekdays.

This preference indicates an opportunity for the hotel to optimize offerings during weekdays.

Data Analysis Queries

#Query7:

What is the highest, lowest, and average lead time for reservations?

```
SELECT MAX(lead_time) AS highest_lead_time,MIN(lead_time) AS lowest_lead_time
FROM hotel_data;
```

highest_lead_time	lowest_lead_time
443	0

```
SELECT AVG(lead_time) AS avg_lead_time
FROM hotel_data;
```

avg_lead_time
83.3000

Insights:

The range of lead times, from 0 to 443 days, suggests diverse booking behaviors among guests. Some prefer last-minute reservations, while others plan well in advance. On average, reservations have a lead time of 83.30 days.

Data Analysis Queries

#Query8:

What is the distribution of market segments for guests making same-day reservations versus those with long lead times (443 days)?

```sql
-- Query for Same-Day Reservations
SELECT market_segment_type, COUNT(market_segment_type) AS total_market_segment
FROM hotel_data
WHERE lead_time = 0
GROUP BY  market_segment_type;
```

market_segment_type	total_market_segment
Online	16
Corporate	7
Offline	2
Complementary	1

```sql
-- Query for Long Lead Time Reservations
SELECT market_segment_type, COUNT(market_segment_type) AS total_market_segment
FROM hotel_data
WHERE lead_time = 443
GROUP BY  market_segment_type;
```

market_segment_type	total_market_segment
Online	1

Insights:

- Urgent last-minute bookings (same-day arrivals) are most common among guests from online, corporate, and offline market segments.
- Guests with the longest lead time predominantly come from the online platform.

Data Analysis Queries

#Query9:

What is the most common market segment type for reservations?

```
SELECT market_segment_type, COUNT(*) AS total_resevations FROM hotel_data
GROUP BY market_segment_type
ORDER BY COUNT(*) DESC
LIMIT 1;
```

market_segment_type	total_resevations
Online	518

Insights:
The online market segment is the most popular choice among guests , showcasing a clear preference for online reservations.

Data Analysis Queries

#Query10:

What is the total number of confirmed reservations, and what percentage of reservations have a "Confirmed" booking status?

```
SELECT COUNT(*) AS confirmed_reservations FROM hotel_data
WHERE booking_status = "Not_Canceled";
```

confirmed_reservations
493

```
-- successful reservation %
SELECT ROUND((SUM(CASE WHEN booking_status = "Not_Canceled" THEN 1 ELSE 0 END)/COUNT(*))*100,2) AS successful_reservation_percent
FROM hotel_data;
```

successful_reservation_percent
70.43

Insights:
Out of 700 reservations, 493 are confirmed, indicating a success rate of approximately 70.43%. This suggests a high rate of successful reservations.

Data Analysis Queries

#Query11:

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_data;
```

total_adults	total_children
1316	69

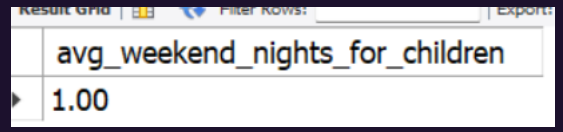
Insights:
The majority of reservations, totaling 1,316, involve adult guests, while a smaller number, 69, include children. This highlights the hotel's primary appeal to adult.

Data Analysis Queries

#Query12:

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

```
SELECT ROUND(AVG(no_of_weekend_nights),2) AS avg_weekend_nights_for_children
FROM hotel_data
WHERE no_of_children > 0;
```



avg_weekend_nights_for_children
1.00

Insights:
On average, reservations with children involve a one-night stay on weekends, highlighting a preference for weekend stays, particularly suitable for families.

Data Analysis Queries

#Query13:

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

```
SELECT MONTH(arrival_date) AS month, MONTHNAME(arrival_date) AS month_name,COUNT(*) AS total_reservations
FROM hotel_data
GROUP BY MONTH(arrival_date) ,MONTHNAME(arrival_date)
ORDER BY COUNT(*)  DESC ;
```

month	month_name	total_reservations
1	January	11
2	February	28
3	March	52
4	April	67
5	May	55
6	June	84
7	July	44
8	August	70
9	September	80
10	October	103
11	November	54
12	December	52

Insights:
October stands out as the peak reservation month followed by June and September . In contrast, January records the lowest number of reservations, indicating a quieter period.

Data Analysis Queries

#Query14:

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

```
SELECT room_type_reserved, ROUND(AVG(no_of_weekend_nights + no_of_week_nights),2) AS avg_total_nights FROM hotel_data
GROUP BY room_type_reserved
ORDER BY AVG(no_of_weekend_nights + no_of_week_nights) DESC ;
```

room_type_reserved	avg_total_nights
Room_Type 4	3.80
Room_Type 6	3.61
Room_Type 2	3.00
Room_Type 1	2.88
Room_Type 7	2.67
Room_Type 5	2.50

Insights:
Guests staying in Room Type 4 tend to spend the most nights on average (3.80), while those in Room Type 5 have the lowest average stay duration (2.50).

Data Analysis Queries

#Query15:

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

```sql
SELECT room_type_reserved, COUNT(*) AS total_reservations , ROUND(AVG(avg_price_per_room),2) AS avg_price_per_room
FROM hotel_data
WHERE no_of_children > 0
GROUP BY room_type_reserved
ORDER BY COUNT(*) DESC
LIMIT 1 ;
```

room_type_reserved	total_reservations	avg_price_per_room
Room_Type 1	24	123.12

Insights:
For reservations involving children, Room Type 1 is the preferred choice, with an average room price of 123.12.

Data Analysis Queries

#Query16:

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

```sql
SELECT market_segment_type , ROUND(AVG(avg_price_per_room),2) AS highest_avg_price_per_room
FROM hotel_data
GROUP BY market_segment_type
ORDER BY highest_avg_price_per_room DESC
LIMIT 1;
```

market_segment_type	highest_avg_price_per_room
Online	112.46

Insights:
Online bookings generate the highest average room price, reaching 112.46.

Key Questions

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 in each 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, lowest, and average lead time for reservations?
8. What is the distribution of market segments for guests making same-day reservations versus those with long lead times (443 days)?
9. What is the most common market segment type for reservations?
10. What is the total number of confirmed reservations, and what percentage of reservations have a "Confirmed" booking status?
11. What is the total number of adults and children across all reservations?
12. What is the average number of weekend nights for reservations involving children?
13. How many reservations were made in each month of the year?
14. What is the average number of nights (both weekend and weekday) spent by guests for each room type?
15. For reservations involving children, what is the most common room type, and what is the average price for that room type?
16. Find the market segment type that generates the highest average price per room?

Overall Insight

- There are 700 reservations in the dataset.
- Meal Plan 1 stands out as the most popular choice among guests.
- Reservations involving children have an average room price of 144.6.
- In 2017, there were 123 reservations. Subsequently, in 2018, there was a notable increase in reservations, indicating positive growth.
- Room Type 1 is the preferred choice for bookings.
- Guests show a preference for weekday night stays (656) over weekends (383).
- Diverse booking behaviors with lead times ranging from 0 to 443 days. On average, guests make reservations approximately 83 days (around 3 months) ahead of their arrival dates.
- Urgent last-minute bookings (same-day arrivals) are most common among guests from online, corporate, and offline market segments.

  Guests with the longest lead time predominantly come from the online platform.
- Online market segment is the most popular choice.
- 493 out of 700 reservations are confirmed, indicating a 70.43% success rate.
- Majority of reservations (1,316) involve adult guests.
- Reservations with children suggest a preference for one-night stays on weekends.
- October is the peak reservation month, while January records the lowest reservations.
- Room Type 4 guests prefer longer night stays (average of 3.80 nights), while Room Type 5 guests opt for shorter durations (average of 2.50 nights).
- For reservations involving children, Room Type 1 is the preferred choice, with an average room price of 123.12.
- Online bookings generate the highest average room price, reaching 112.46.

Recommendation

- Implement targeted promotions for Room Type 1 to capitalize on its popularity.
- Tailor marketing strategies to attract online bookings, the most prevalent segment.
- Explore partnerships or promotions to boost reservations during quieter months like January.
- Enhance confirmation and booking processes to maintain the high success rate of reservations.
- Introduce special packages and incentives to attract guests seeking longer stays, especially those choosing Room Type 4.
- Focus on enhancing services and promotions during weekdays to meet the strong demand for reservations on weekday nights (656), creating an opportunity to attract a larger number of guests.
- Offer special discounts or promotions for guests making last-minute bookings. This strategy not only attracts more bookings but also minimizes the chance of cancellations.
- Improve the hotel experience for families by introducing special amenities and activities. Create attractive packages for one-night stays on weekends, designed for families looking for a weekend getaway.
- Continue monitoring and adapting strategies based on changing guest preferences and market trends.

Thank You