

OYO HOTEL BOOKING ANALYSIS



SQL Data Analytics Project Report



Presented by: Rohan Nayak

Date: 19/11/25

OYO Hotel Booking Analysis Using SQL

Prepared By: Rohan Nayak

Role: Aspiring Data Analyst

Tools: MySQL, MySQL Workbench, GitHub

1. Executive Summary

This project focuses on analyzing OYO hotel booking data using **pure SQL (MySQL)** to derive actionable business insights such as booking trends, cancellation rates, revenue distribution, and customer behavior across cities. The objective was to strengthen SQL logic building and problem-solving skills by working on real-world data.

2. Project Objectives

- Analyze hotel booking patterns across cities
- Identify cancellation trends and booking status distribution
- Understand monthly booking behavior (Jan, Feb, Mar)
- Compute revenue and discount impact
- Identify top cities and loyal customers
- Improve SQL logic and query optimization skills

3. Dataset Description

Table: oyo_sales

Column	Description
booking_id	Unique booking ID
customer_id	Unique customer ID
status	Booking status
check_in	Check-in date
check_out	Check-out date
no_of_rooms	Number of rooms booked
hotel_id	Hotel unique ID
amount	Total booking amount

Column	Description
discount	Discount applied
date_of_booking	Booking creation date

Table: oyo_city

Column Description

hotel_id Hotel ID

city Hotel city

4. Data Cleaning & Preprocessing

The following preprocessing steps were performed:

- Fixed encoding errors in column names ('booking_id', 'hotel_id').
- Converted date fields into proper MySQL date format.
- Handled NULL and zero values to avoid calculation errors.
- Verified relationships between booking and hotel datasets.

5. Methodology

1. Import datasets into MySQL
2. Validate schema and data types
3. Perform data cleaning and corrections
4. Execute SQL queries to answer business questions
5. Analyze results and derive insights
6. Document findings with screenshots

6. Business Questions Analyzed

1. Total bookings, hotels, and cities
2. Hotels distribution across cities
3. Average room rate per city
4. Cancellation rate by city
5. Monthly bookings (Jan, Feb, Mar)
6. Revenue and discount analysis

7. Booking status distribution
8. Top 5 cities by bookings
9. Top 10 customers by booking count
10. Single vs multiple room bookings
11. Average length of stay
12. Average discount percentage

7. Analysis & Screenshots

7.1 Database Tables Overview

Tables structure (oyo_sales and oyo_city)

101 • select * from oyo_city, oyo_sales;

102

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

	hotel_id	city	booking_id	customer_id	status	check_in	check_out	no_of_rooms	hotel_id	amount	discount	date_of_booking
▶	998	Hyderabad	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	997	Hyderabad	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	996	Hyderabad	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	995	Pune	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	984	Hyderabad	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	979	Hyderabad	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	978	Chennai	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022
	975	Bangalore	170	9197	Cancelled	14-01-2022	15-01-2022	1	757	3137	796	01-01-2022

Result 18 ×

7.2 Sample SQL Query Execution

Example: Average Room Rate Per City

```
15      -- 3. Average room rates of different cities
16 •  SELECT c.city,
17    ROUND(AVG((s.amount - s.discount) / NULLIF(s.no_of_rooms,0)),2) AS avg_rate_per_room
18    FROM oyo_sales s
19    JOIN oyo_city c ON s.hotel_id = c.hotel_id
20    GROUP BY c.city;
21
```

Result Grid | Filter Rows: Export: Wrap Cell Content: Fetch rows:

city	avg_rate_per_room
Bangalore	3339.29
Gurgaon	2079.19
Mumbai	5715.58
Delhi	3213.04
Hyderabad	3225.16
Jaipur	2260.72
Noida	2347.81
Chennai	2999.50

Result 20 ×

8. Key Insights

- Certain cities show much higher booking volume.
- Cancellation patterns vary significantly between locations.
- Multi-room bookings generate higher revenue.
- Discounts influence booking frequency but affect profitability.

9. Challenges Faced

- Handling incorrect encoding in CSV/Excel imports
- Managing inconsistent date formats
- Avoiding division by zero in calculations
- Query optimization for large datasets

10. Conclusion

This project provided valuable hands-on experience in SQL-driven data analysis. It helped improve logical thinking, query writing efficiency, and interpreting business data for insights.

12. GitHub Repository

 <https://github.com/rohannayak360/OYO-Hotel-Booking-Analysis.git>