

Understand the data

Data contains all the meta information regarding the columns described in the CSV files. we have provided 5 CSV files:

1. dim_date
2. dim_hotels
3. dim_rooms
4. fact_aggregated_bookings
5. fact_bookings

Column Description for dim_date:

1. date: This column represents the dates present in May, June and July.
2. mmm yy: This column represents the date in the format of mmm yy (monthname year).
3. week no: This column represents the unique week number for that particular date.
4. day_type: This column represents whether the given day is Weekend or Weekday.

Column Description for dim_hotels:

1. property_id: This column represents the Unique ID for each of the hotels.
2. property_name: This column represents the name of each hotel.
3. category: This column determines which class[Luxury, Business] a particular hotel/property belongs to.
4. city: This column represents where the particular hotel/property resides in.

Column Description for dim_rooms:

1. room_id: This column represents the type of room[RT1, RT2, RT3, RT4] in a hotel.
2. room_class: This column represents to which class[Standard, Elite, Premium, Presidential] particular room type belongs.

Column Description for fact_aggregated_bookings:

1. property_id: This column represents the Unique ID for each of the hotels.
2. check_in_date: This column represents all the check_in_dates of the customers.
3. room_category: This column represents the type of room[RT1, RT2, RT3, RT4] in a hotel.
4. successful_bookings: This column represents all the successful room bookings that happen for a particular room type in that hotel on that particular date.
5. capacity: This column represents the maximum count of rooms available for a particular room type in that hotel on that particular date.

Column Description for fact_bookings:

1. booking_id: This column represents the Unique Booking ID for each customer when they booked their rooms.
2. property_id: This column represents the Unique ID for each of the hotels
3. booking_date: This column represents the date on which the customer booked their rooms.
4. check_in_date: This column represents the date on which the customer checkin(entered) at the hotel.
5. check_out_date: This column represents the date on which the customer checkout(left) of the hotel.
6. no_guests: This column represents the number of guests who stayed in a particular room in that hotel.
7. room_category: This column represents the type of room[RT1, RT2, RT3, RT4] in a hotel.
8. booking_platform: This column represents in which way the customer booked his room.
9. ratings_given: This column represents the ratings given by the customer for hotel services.
10. booking_status: This column represents whether the customer cancelled his booking[Cancelled], successfully stayed in the hotel[Checked Out] or booked his room but not stayed in the hotel[No show].
11. revenue_generated: This column represents the amount of money generated by the hotel from a particular customer.
12. revenue_realized: This column represents the final amount of money that goes to the hotel based on booking status. If the booking status is cancelled, then 40% of the revenue generated is deducted and

the remaining is refunded to the customer. If the booking status is Checked Out/No show, then full revenue generated will go to hotels