Data Analysis Hackathon: Hospitality Data Analysis

Project: Provide Insights to Revenue Team in Hospitality Domain

GDS Grands owns multiple five-star hotels across India. They have been in the hospitality industry for the past 20 years. Due to strategic moves from other competitors and ineffective decision-making in management, GDS Grands are losing its market share and revenue in the luxury/business hotels category.

As a strategic move, the managing director of GDS Grands wanted to incorporate "Business and Data Intelligence" in order to regain their market share and revenue. However, they do not have an in-house data analytics team to provide them with these insights.

Their revenue management team had decided to hire a 3rd party service provider to provide them insights from their historical data.

Task:

Mr. Analyst is the data analyst. He has been briefed about the task in the stakeholder business review meeting. Now Imagine yourself as Mr. Analyst and play the role of the new data analyst who is excited to build this dashboard and perform the following task:

- Create the metrics according to the metric list.
- Create a dashboard according to the mock-up provided by stakeholders.
- Create relevant insights that are not provided in the metric list/mock-up dashboard.

Metrics List:

Sno	Measures	Description
1	Revenue	Sum of revenue_realized
2	Total Bookings	Count of booking_id in fact_bookings
3	Average Rating	Average of ratings_given
4	Total Capacity	Sum of capacity
5	Total Succesful bookings	Sum of successful_bookings from fact_bookings
6	Occupancy %	Ratio of Total Successful Bookings to Total Capacity
7	Total Cancelled Bookings	Count of booking_id in which booking_status = "Cancelled"
8	Cancellation Rate	Ratio of 'Total Cancelled Bookings' to 'Total Bookings'

Meta Data:

We have provided 5 files(folder attached):

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 Weekeday.

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 check-in(entered) at the hotel.

5. check_out_date: This column represents the date on which the customer check-out(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 goes to hotels.

Evaluation Criteria:

● Completeness of the Project: End-to-end functionality from source to destination

● Data Processing Efficiency: Speed and precision of calculations.

● Dashboard Usability: Clarity, real-time capabilities, and user experience.

● Analysis: Key findings and insights drawn from the data

● Documentation: Clear documentation of architecture, data flows, tools/components used and outcomes.

Must do things for project submission:

● Complete code should be checked in into GitHub along with readme file, architecture

diagram, installation steps and step by step process to execute the project

● Teams need to create a 4-5 minutes of working Video demo of their project, need to

explain the architecture and working of complete project with data and present the final insights.

● No project submission will be accepted if above 2 points are not taken care