Data Analyst Project

OLA Data Analyst Project

Project Constraints:

Make sure orders cancelled by customers should not be more than 7%
Make sure orders cancelled drivers should not be more than 18%

Also, increase the number of orders on weekends and match days. Keep match day by using the following dates.
keep incomplete rides less than 6%
Keep order value high on weekends
Increase bookings on weekends and special event days.

OLA Data Analyst Project

SQL Questions:

1. Retrieve all successful bookings:
2. Find the average ride distance for each vehicle type:
3. Get the total number of cancelled rides by customers:
4. List the top 5 customers who booked the highest number of rides: 5. Get the number of rides cancelled by drivers due to personal and car-related issues: 6. Find the maximum and minimum driver ratings for Prime Sedan bookings: 7. Retrieve all rides where payment was made using UPI:
8. Find the average customer rating per vehicle type:
9. Calculate the total booking value of rides completed successfully:
10. List all incomplete rides along with the reason:

Power BI Questions:

1. Ride Volume Over Time
2. Booking Status Breakdown
3. Top 5 Vehicle Types by Ride Distance
4. Average Customer Ratings by Vehicle Type
5. cancelled Rides Reasons
6. Revenue by Payment Method
7. Top 5 Customers by Total Booking Value
8. Ride Distance Distribution Per Day
9. Driver Ratings Distribution
10. Customer vs. Driver Ratings

Data Columns

1. Date
2. Time
3. Booking_ID 4. Booking_Status 5. Customer_ID 6. Vehicle_Type 7. Pickup_Location 8. Drop_Location 9. V_TAT
10. C_TAT
11. cancelled_Rides_by_Customer 12. cancelled_Rides_by_Driver 13. Incomplete_Rides
14. Incomplete_Rides_Reason 15. Booking_Value
16. Payment_Method
17. Ride_Distance
18. Driver_Ratings
19. Customer_Rating

OLA Data Analyst Project

SQL Answers:

1. Retrieve all successful bookings:

```
SELECT * FROM bookings WHERE Booking_Status = 'Success';
```

2. Find the average ride distance for each vehicle type:

```
SELECT Vehicle_Type, AVG(Ride_Distance) as avg_distance FROM bookings GROUP BY Vehicle_Type;
```

3. Get the total number of cancelled rides by customers:

```
SELECT COUNT(*) FROM bookings WHERE Booking_Status = 'cancelled by Customer';
```

4. List the top 5 customers who booked the highest number of rides:

```
SELECT Customer_ID, COUNT(Booking_ID) as total_rides FROM bookings GROUP BY Customer_ID ORDER BY total_rides DESC LIMIT 5;
```

5. Get the number of rides cancelled by drivers due to personal and car-related issues:

```
SELECT COUNT(*) FROM bookings WHERE cancelled_Rides_by_Driver = 'Personal & Car related issue';
```

6. Find the maximum and minimum driver ratings for Prime Sedan bookings:

```
SELECT MAX(Driver_Ratings) as max_rating, MIN(Driver_Ratings) as min_rating FROM bookings WHERE Vehicle_Type = 'Prime Sedan';
```

7. Retrieve all rides where payment was made using UPI:

```
SELECT * FROM bookings WHERE Payment_Method = 'UPI';
```

8. Find the average customer rating per vehicle type:
SELECT Vehicle_Type, AVG(Customer_Rating) as avg_customer_rating FROM bookings GROUP BY Vehicle_Type;

9. Calculate the total booking value of rides completed successfully: SELECT SUM(Booking_Value) as total_successful_value FROM bookings WHERE Booking_Status = 'Success';

10. List all incomplete rides along with the reason:
SELECT Booking_ID, Incomplete_Rides_Reason FROM bookings WHERE Incomplete_Rides = 'Yes';

OLA Data Analyst Project

Power BI Answers:

Segregation of the views:

1. Overall
- Ride Volume Over Time
- Booking Status Breakdown

2. Vehicle Type
- Top 5 Vehicle Types by Ride Distance

3. Revenue
- Revenue by Payment Method
- Top 5 Customers by Total Booking Value
- Ride Distance Distribution Per Day

4. Cancellation
- Cancelled Rides Reasons (Customer)
- cancelled Rides Reasons(Drivers)

5. Ratings
- Driver Ratings
- Customer Ratings

Answers:

1. Ride Volume Over Time: A time-series chart showing the number of rides per day/week. **2. Booking Status Breakdown:** A pie or doughnut chart displaying the proportion of different booking statuses (success, cancelled by the customer, cancelled by the driver, etc.). **3. Top 5 Vehicle Types by Ride Distance:** A bar chart ranking vehicle types based on the total distance

covered.
4. Average Customer Ratings by Vehicle Type: A column chart showing the average customer ratings for different vehicle types.
5. cancelled Rides Reasons: A bar chart that highlights the common reasons for ride cancellations by customers and drivers.
6. Revenue by Payment Method: A stacked bar chart displaying total revenue based on payment methods (Cash, UPI, Credit Card, etc.).
7. Top 5 Customers by Total Booking Value: A leaderboard visual listing customers who have spent the most on bookings.
8. Ride Distance Distribution Per Day: A histogram or scatter plot showing the distribution of ride distances for different Dates.
9. Driver Rating Distribution: A box plot visualizing the spread of driver ratings for different vehicle types.
10. Customer vs. Driver Ratings: A scatter plot comparing customer and driver ratings for each completed ride, analyzing correlations.

OLA Data Analyst Project

SQL Questions & Answers

```
Create Database Ola;
Use Ola;
```

#1. Retrieve all successful bookings:

```
Create View Successful_Bookings As
SELECT * FROM bookings
WHERE Booking_Status = 'Success';
```

#2. Find the average ride distance for each vehicle type:

```
Create View ride_distance_for_each_vehicle As
SELECT Vehicle_Type, AVG(Ride_Distance)
as avg_distance FROM bookings
GROUP BY Vehicle_Type;
```

#3. Get the total number of cancelled rides by customers:

```
Create View cancelled_rides_by_customers As
SELECT COUNT(*) FROM bookings
WHERE Booking_Status = 'cancelled by Customer';
```

#4. List the top 5 customers who booked the highest number of rides:

```
Create View Top_5_Customers As
SELECT Customer_ID, COUNT(Booking_ID) as total_rides
FROM bookings
GROUP BY Customer_ID
ORDER BY total_rides DESC LIMIT 5;
```

#5. Get the number of rides cancelled by drivers due to personal and car-related issues:

```
Create View Rides_cancelled_by_Drivers_P_C_Issues As
SELECT COUNT(*) FROM bookings
WHERE cancelled_Rides_by_Driver = 'Personal & Car related issue';
```

#6. Find the maximum and minimum driver ratings for Prime Sedan bookings:

```
Create View Max_Min_Driver_Rating As
SELECT MAX(Driver_Ratings) as max_rating,
MIN(Driver_Ratings) as min_rating
FROM bookings WHERE Vehicle_Type = 'Prime Sedan';
```

OLA Data Analyst Project

#7. Retrieve all rides where payment was made using UPI:

```
Create View UPI_Payment As
SELECT * FROM bookings
WHERE Payment_Method = 'UPI';
```

#8. Find the average customer rating per vehicle type:

```
Create View AVG_Cust_Rating As
SELECT Vehicle_Type, AVG(Customer_Rating) as avg_customer_rating
FROM bookings
GROUP BY Vehicle_Type;
```

#9. Calculate the total booking value of rides completed successfully:

```
Create View total_successful_ride_value As
SELECT SUM(Booking_Value) as total_successful_ride_value
FROM bookings
WHERE Booking_Status = 'Success';
```

#10. List all incomplete rides along with the reason:

```
Create View Incomplete_Rides_Reason As
```

```
SELECT Booking_ID, Incomplete_Rides_Reason
FROM bookings
WHERE Incomplete_Rides = 'Yes';
```

Retrieve All Answers

#1. Retrieve all successful bookings:
Select * From Successful_Bookings;

#2. Find the average ride distance for each vehicle type:
Select * from ride_distance_for_each_vehicle;

#3. Get the total number of cancelled rides by customers:
Select * from cancelled_rides_by_customers;

#4. List the top 5 customers who booked the highest number of rides:
Select * from Top_5_Customers;

OLA Data Analyst Project

#5. Get the number of rides cancelled by drivers due to personal and car-related issues:
Select * from Rides_cancelled_by_Drivers_P_C_Issues;

#6. Find the maximum and minimum driver ratings for Prime Sedan bookings:
Select * from Max_Min_Driver_Rating;

#7. Retrieve all rides where payment was made using UPI:
Select * from UPI_Payment;

#8. Find the average customer rating per vehicle type:
Select * from AVG_Cust_Rating;

#9. Calculate the total booking value of rides completed successfully:
Select * from total_successful_ride_value;

#10. List all incomplete rides along with the reason:
Select * from Incomplete_Rides_Reason;