



# Data Analyst Project

BY BISWAJIT SAMAL

# 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 canceled rides by customers:
4. List the top 5 customers who booked the highest number of rides:
5. Get the number of rides canceled 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. Canceled 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            | 10. C_TAT                      |
| 2. Time            | 11. Canceled_Rides_by_Customer |
| 3. Booking_ID      | 12. Canceled_Rides_by_Driver   |
| 4. Booking_Status  | 13. Incomplete_Rides           |
| 5. Customer_ID     | 14. Incomplete_Rides_Reason    |
| 6. Vehicle_Type    | 15. Booking_Value              |
| 7. Pickup_Location | 16. Payment_Method             |
| 8. Drop_Location   | 17. Ride_Distance              |
| 9. V_TAT           | 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 canceled rides by customers:

```
SELECT COUNT(*) FROM bookings WHERE Booking_Status = 'Canceled 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 canceled by drivers due to personal and car-related issues:

```
SELECT COUNT(*) FROM bookings WHERE Canceled_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**
  - Canceled Rides Reasons (Customer)
  - Canceled 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 donut chart displaying the proportion of different booking statuses (success, canceled by customer, canceled by 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. Canceled 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 Ratings 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 canceled rides by customers:

Create View canceled\_rides\_by\_customers As  
SELECT COUNT(\*) FROM bookings  
WHERE Booking\_Status = 'Canceled 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 canceled by drivers due to personal and car-related issues:

Create View Rides\_Canceled\_by\_Drivers\_P\_C\_Issues As  
SELECT COUNT(\*) FROM bookings  
WHERE Canceled\_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 canceled rides by customers:**

```
Select * from canceled_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 canceled by drivers due to personal and car-related issues:**

Select \* from Rides\_Canceled\_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;