

# OLA Data Analyst Project

## ChatGPT Prompt to Create Data

**Please create a spreadsheet with 1 lac rows, for Bengaluru city. Give the following columns. The data will be for 1 month. use the following column.**

1. Date
2. Time
3. Booking ID
4. Booking Status
5. Customer ID
6. Vehicle Type:-
  - Auto
  - Prime Plus
  - Prime Sedan
  - Mini
  - Bike
  - e-Bike
  - Prime SUV
7. Pickup Location (Create dummy location points Take any 50 areas from Bangalore)
8. Drop Location (Take from dummy pickup locations)
9. Avg VTAT (Time taken to arrive at the vehicle)
10. Avg CTAT (Time taken to arrive the Customer)
11. Cancelled Rides by Customer (0/1)
12. Reason for cancelling by Customer:-
  - Driver is not moving towards pickup location.
  - Driver asked to cancel.
  - AC is not working (Only for 4-wheelers).
  - Change of plans
  - Wrong Address.
13. Cancelled Rides by Driver (0/1)
14. Reason for Cancelling by Driver
  - Personal & Car related issues

- Customer related issue
- The customer was coughing/sick
- More than permitted people in there

15.Incomplete Rides (0/1)

16.Incomplete Rides Reason

- Customer Demand
- Vehicle Breakdown
- Other Issue

17.Payment Type

- Cash
- Card
- UPI

18.Booking Value

19.Ride Distance

20.Driver Ratings

21.Customer Rating

Keep the overall booking status success for this data at 62%. If the booking status is successful, then only fare charge ratings, average VTAT, average CTAT, and other data will be there.

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

# 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. List average Driver & Customer ratings of all 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. 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

# OLA Data Analyst Project

## SQL Answers

### 1. Retrieve all Successful bookings:

```
select * from OLA  
where Booking_Status = 'Successful'
```

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

```
select  
Vehicle_Type, round(avg(Ride_Distance_km),2) as Avg_Ride_Distance  
from OLA  
group by Vehicle_Type  
order by Avg_Ride_Distance
```

### 3. Get the total no of rides cancelled by customers

```
select  
count(*) as No_of_Rides_Cancelled_by_Customer  
from OLA  
where Booking_Status = 'Cancelled by customer'
```

### 4. List the top 5 customers who booked the highest no of rides

```
select Top 5  
Customer_ID, count(Booking_ID) as Total_Rides  
from OLA  
group by Customer_ID  
order by Total_Rides desc
```

### 5. Get the no of rides cancelled by drivers due to personal and car-related issues

```
select  
count(Reason_for_Cancelling_by_Driver) as Count  
from OLA  
where Reason_for_cancelling_by_Driver = 'Personal & Car related issues'
```

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

```
select
Vehicle_Type,
min(Driver_Ratings) as Min_Driver_Ratings,
max(Driver_Ratings) as Max_Driver_Ratings
from OLA
where Vehicle_Type = 'Prime Sedan'
group by Vehicle_Type
order by Vehicle_Type
```

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

```
select *
from OLA
where Payment_Type = 'UPI'
```

**8. List average Driver & Customer ratings of all Vehicle type**

```
select Vehicle_Type,
round(avg(Driver_Ratings),2) as Avg_Driver_Ratings,
round(avg(Customer_Ratings),2) as Avg_Customer_Ratings
from OLA
group by Vehicle_Type
```

**9. Calculate the total booking value of rides completed safely**

```
select
Booking_Status,
round(sum(Booking_Value),2) as Total_Booking_Value
from OLA
where Booking_Status = 'Successful'
group by Booking_Status
order by Total_Booking_Value
```

**10. List all incomplete rides along with the reason**

```
select *
from OLA
where Booking_Status = 'Incomplete'
```

# OLA Data Analyst Project

## Power BI Answers

### Segregation of the views:

#### 1. Overall

- Ride Volume Over Time
- Booking Status Breakdown
- Top 5 Vehicle Types by Ride Distance

#### 2. Vehicle Type

- 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 by Customers
- Cancelled Rides by Drivers

#### 5. Ratings

- Driver Ratings
- Customer Ratings