

Data Analyst Project

By Alhan Siddiqui

# 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

# SQL Answers:

### Retrieve all successful bookings:

SELECT \* FROM bookings WHERE Booking\_Status = 'Success';

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

SELECT Vehicle\_Type, AVG(Ride\_Distance) as avg\_distance FROM bookings GROUP BY Vehicle\_Type;

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

SELECT COUNT(\*) FROM bookings WHERE Booking\_Status = 'cancelled by Customer';

### 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;

1. **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';
2. **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';

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

SELECT \* FROM bookings WHERE Payment\_Method = 'UPI';

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

SELECT Vehicle\_Type, AVG(Customer\_Rating) as avg\_customer\_rating FROM bookings GROUP BY Vehicle\_Type;

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

SELECT SUM(Booking\_Value) as total\_successful\_value FROM bookings WHERE Booking\_Status = 'Success';

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

SELECT Booking\_ID, Incomplete\_Rides\_Reason FROM bookings WHERE Incomplete\_Rides = 'Yes';

# Power BI Answers:

Segregation of the views:

### Overall

* + Ride Volume Over Time
  + Booking Status Breakdown

1. **Vehicle Type**
   * Top 5 Vehicle Types by Ride Distance
2. **Revenue**
   * Revenue by Payment Method
   * Top 5 Customers by Total Booking Value
   * Ride Distance Distribution Per Day
3. **Cancellation**
   * Cancelled Rides Reasons (Customer)
   * cancelled Rides Reasons(Drivers)
4. **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.

**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';

### #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;

### #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;