



# Data Analyst Project

## By Ritesh Kumar

### Rapido Data Analyst Project

Check My Profile:

My GitHub – <https://github.com/ritessinghh>

My LinkedIn – [linkedin.com/in/riteshhsingh](https://www.linkedin.com/in/riteshhsingh)

My Gmail – [ritessingh108@gmail.com](mailto:ritessingh108@gmail.com)

## 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:

## SQL Answer

-- 1. Retrieve all successful bookings:

```
select * from booking
where booking_status = 'Success';
```

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

```
select avg(ride_distance), vehicle_type from booking
group by vehicle_type;
```

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

```
select count(booking_status) from booking
where booking_status = 'Canceled by Driver';
```

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

```
select count(booking_id) as total_ride, customer_id from booking
group by customer_id
order by total_ride
limit 5;
```

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

```
select count(canceled_rides_by_driver) from booking
where canceled_rides_by_driver = 'Personal & Car related issues';
```

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

```
select max(driver_ratings), min(driver_ratings) from booking
where vehicle_type = 'Prime Sedan';
```

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

```
select * from booking
where payment_method = 'UPI';
```

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

```
select vehicle_type, avg(customer_rating) as Avg_customer_Rating
from booking
group by vehicle_type
order by Avg_customer_Rating desc;
```

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

```
select sum(booking_value) as Total_Booking_Value, booking_status
from booking
where booking_status = 'Success';
```

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

```
select incomplete_rides, incomplete_rides_reason from booking
where incomplete_rides = 'Yes';
```

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

## 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.