OLA Data Analyst Project

- ✓ 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 in Food Category
- ✓ keep around 67 Indian keep order ID with 10 digits starting with CNR and then digits keep orders under 500 value 70% keep orders above 500 value 28% keep remaining orders above 1000

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

```
Create Database ola;

use ola;

SHOW TABLES
```

1. Retriving all successfull Bookings

```
CREATE VIEW Successfull_Bookings AS

SELECT * FROM  bookings

WHERE Booking_Status = 'Success' ;


SELECT * FROM  Successfull_Bookings
```

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

SELECT * FROM RIDE_DISTANCE_FOR_EACH_VEHICLE
```

3. GET THE TOTAL NUMBER OF CANCELED RIDES BY CUSTOMERS:

```
CREATE VIEW Canceled_Ride_By_Customers AS

SELECT COUNT(*) FROM bookings

WHERE Booking_Status = 'Canceled by Customer'

SELECT * FROM Canceled_Ride_By_Customers
```

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

```
SELECT * FROM  bookings

CREATE VIEW top_5_customers_who_booked_the_highest_number_of_rides AS

SELECT Customer_ID , COUNT(Booking_ID) AS total_rides

FROM bookings

GROUP BY Customer_ID

ORDER BY total_rides DESC LIMIT 5;

SELECT * FROM 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 :

```
SELECT * FROM  bookings

CREATE VIEW rides_canceled_by_drivers_due_to_p_c_issues  AS

SELECT COUNT(*) FROM bookings

WHERE Canceled_Rides_by_Driver = 'Personal & Car related issue'

SELECT * FROM rides_canceled_by_drivers_due_to_p_c_issues
```

6 FIND THE MAXIMUM AND MINIMUM DRIVER RATINGS FOR PRICE SEDAN BOOKINGS :

```
SELECT * FROM  bookings

CREATE VIEW MAX_AND_MIN_DRIVER_RATINGS_FOR_PRICE_SEDAN AS

SELECT

  MAX(Driver_Ratings) AS max_rating ,

  MIN(Driver_Ratings)AS min_rating

FROM bookings

WHERE Vehicle_Type = 'Prime Sedan'

SELECT* FROM MAX_AND_MIN_DRIVER_RATINGS_FOR_PRICE_SEDAN
```

7. RETRIVE ALL RIDES WHERE PAYMENT WAS MADE USING UPI:

```
CREATE VIEW UPI_PAYMENT AS

SELECT * FROM  bookings

WHERE Payment_Method = 'UPI'

SELECT * FROM UPI_PAYMENT
```

8. FIND THE AVERAGE CUSTOMER RATING PER VEHICLE TYPE:

```
CREATE VIEW AVG_CUSTOMER_RATING AS

SELECT Vehicle_Type,AVG(Customer_Rating) AS avg_customer_rating

FROM bookings

GROUP BY Vehicle_Type

SELECT * FROM AVG_CUSTOMER_RATING
```

9. CALCULATE THE TOTAL BOOKING VALUE OF RIDES COMPLETED SUCCESSFULLY :

```
SELECT * FROM  bookings
```

```
CREATE VIEW total_successfull_ride_value AS
SELECT SUM(Booking_Value) AS total_successfull_ride_value
FROM bookings
WHERE Booking_Status = 'Success'

SELECT * FROM total_successfull_ride_value
```

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'
SELECT * FROM Incomplete_Rides_Reason
```

```
### 1. Retriving all successfull Bookings
SELECT * FROM  Successfull_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_Ride_By_Customers
#### 4. List THE top 5 customers who booked the highest number of rides:
SELECT * FROM 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 :
SELECT * FROM rides_canceled_by_drivers_due_to_p_c_issues
#### 6 FIND THE MAXIMUM AND MINIMUM DRIVER RATINGS FOR PRICE SEDAN BOOKINGS :
SELECT* FROM MAX_AND_MIN_DRIVER_RATINGS_FOR_PRICE_SEDAN
#### 7. RETRIVE ALL RIDES WHERE PAYMENT WAS MADE USING UPI:
SELECT * FROM UPI_PAYMENT
#### 8. FIND THE AVERAGE CUSTOMER RATING PER VEHICLE TYPE:
SELECT * FROM AVG_CUSTOMER_RATING
##### 9. CALCULATE THE TOTAL BOOKING VALUE OF RIDES COMPLETED SUCCESSFULLY :
SELECT * FROM total_successfull_ride_value
```

10. LIST ALL INCOMPLETE RIDES ALONG WITH THE REASON :

SELECT * FROM Incomplete_Rides_Reason

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.