**Understanding the Lyft Dataset:**

1. Date- Date of Booking

2. Time- Time of Booking

3. Booking ID – Unique for each booking

4. Booking Status – 1.Success, Failure (subdivided as 2.Driver not found, 3.Cancelled by Customer, 4.Cancelled by Driver). The Status is one of the four values.

5. Customer ID – Unique to each customer but one CustomerID can have many bookings in a month.

6. Vehicle Type - Auto - Prime Plus - Prime Sedan - Mini - Bike - eBike - Prime SUV

7. Pickup Location (dummy location points Take any 50 areas from Bangalore,India)

8. Drop Location (dummy locations points Take any 50 areas from Bangalore,India)

9. Avg VTAT (Time taken to arrive at the vehicle)

10. Avg CTAT (Time taken to arrive to the Customer)

11. Cancelled Rides by Customer – contains any of the reasons given by customer for cancelling the ride:  
 a. Change of Plans

b. Driver asked to cancel

c. Driver not moving to pickup location

d. Wrong address entered

e. Null if the cancellation was not done by customer

12. Cancelled Rides by Driver - contains any of the reasons given by driver for cancelling the ride:  
 a. Personal & Car related issues

b. Customer related issue

c. The customer was coughing/sick

d. More than permitted people in there

13. Incomplete Rides – Contains yes/no (if the booking was a success), null (if the booking is a failure)

14. Incomplete Rides Reason – If the previous column is yes, then this column gives the reason for incomplete ride from one of the following.

1. Customer Demand
2. Vehicle Breakdown
3. Other Issue

15. Booking Value – Amount to be paid in rupees, (calculated internally using formula)

16. Payment Method – UPI, Debit card, credit card, cash, if ride is unsuccessful- null

17. Ride Distance – Distance covered during the ride

18. Driver Ratings – given to driver by customer

19. Customer Rating – given to customer by driver

Other Requirements:  
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 apply.

orders cancelled by customers should not be more than 7%

orders cancelled drivers should not be more than 18%

keep incomplete rides less than 6%

For a given booking either Cancelled Rides by Customer or Cancelled Rides by Driver (strictly one can only have valid reason, both can have null but both cannot have reasons).

**Note**: The dataset is synthetically generated using AI tools to apply Power Bi and Data Analytics knowledge for similar cab ride sharing services.

**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