### ****SQL / POWER BI PROJECT QUESTIONS****

### ****Ride Booking Insights****

1. **What is the total number of bookings?**
2. **What is the percentage distribution of all ride statuses?**
3. **How many average bookings were made per day?**
4. **Which day of the week has the highest number of bookings?**

### ****Customer Insights****

1. **Who are the top 5 customers based on the highest number of bookings?**
2. **Who are the top 5 customers based on the highest total amount paid?**

### ****Vehicle and Ratings Analysis****

1. **Which vehicle type is booked the most?**
2. **Which vehicle type has the highest average driver rating?**

### ****Cancellation & Incomplete Ride Analysis****

1. **What are the most common reasons for ride cancellations by customers?**
2. **What is the most common reason for ride cancellations by drivers?**
3. **What is the most common reason for incomplete rides?**

### ****Location & Demand Analysis****

1. **Which pickup location has the highest number of bookings?**
2. **Which drop location is the most common?**

### ****Time-Based Analysis****

1. **What are the peak hours for ride bookings?**
2. **How does the number of bookings vary by weekday vs. weekend?**

**Power BI Questions**

1. **Overall -**

* Ride volume over time
* Booking status breakdown

1. **Vehicle Type -**

* Top 5 vehicle types by ride distance

1. **Revenue -**

* Revenue distribution of vehicle types
* Top 5 customers by total booking value
* Rides distribution per day

1. **Cancellation -**

* Canceled rides reasons (Customer)
* Canceled rides reasons (Driver)

1. **Ratings -**

* Driver ratings
* Customer ratings

### ****SQL / POWER BI PROJECT ANSWERS****

### ****Ride Booking Insights****

1. **What is the total number of bookings?**

SELECT COUNT (\*) AS TOTAL\_NO\_OF\_BOOKINGS -- Count (\*) gives the number of total bookings made

FROM FINAL\_BOOKINGS;

1. **What is the percentage distribution of all ride statuses?**

SELECT

CASE -- it is like conditional statements or if/else in python

WHEN BOOKING\_STATUS IN ('Cancelled by Customer', 'Cancelled by Driver', 'Success', 'Incomplete')

THEN BOOKING\_STATUS

ELSE 'Others'

END AS BOOKING\_STATUS\_GROUPED,

COUNT (\*) AS TOTAL\_RIDES,

ROUND (((COUNT (\*) \* 100/ (SELECT COUNT (\*) FROM FINAL\_BOOKINGS))),2) || '%' AS AVERAGE\_PERCENTAGE

FROM FINAL\_BOOKINGS

GROUP BY BOOKING\_STATUS\_GROUPED

ORDER BY TOTAL\_RIDES DESC ;

1. **How many average bookings were made per day?**

SELECT ROUND(AVG(RIDES\_BOOKED)) AS AVERAGE\_BOOKINGS\_MADE\_PER\_DAY

FROM (

SELECT COUNT (\*) AS RIDES\_BOOKED, DATE

FROM FINAL\_BOOKINGS

GROUP BY DATE

ORDER BY DATE ASC) AS DAILY\_BOOKINGS;

1. **Which day of the week has the highest number of bookings?**

SELECT EXTRACT(ISODOW FROM DATE) AS WEEKDAY, TO\_CHAR(DATE, 'DAY') AS WEEKDAY\_NAME, COUNT(\*) AS TOTAL\_BOOKINGS

FROM FINAL\_BOOKINGS

GROUP BY WEEKDAY, WEEKDAY\_NAME

ORDER BY WEEKDAY DESC

LIMIT 1;

### ****Customer Insights****

1. **Who are the top 5 customers based on the highest number of bookings?**

SELECT COUNT (\*) AS TOTAL\_BOOKINGS, CUSTOMER\_ID

FROM FINAL\_BOOKINGS

GROUP BY CUSTOMER\_ID

ORDER BY TOTAL\_BOOKINGS DESC

LIMIT 5;

1. **Who are the top 5 customers based on the highest total amount paid?**

SELECT

CUSTOMER\_ID,

SUM(COALESCE(BOOKING\_VALUE, 0)) AS TOTAL\_AMOUNT\_PAID

FROM FINAL\_BOOKINGS

GROUP BY CUSTOMER\_ID

ORDER BY TOTAL\_AMOUNT\_PAID DESC

LIMIT 5;

### ****Vehicle and Ratings Analysis****

1. **Which vehicle type is booked the most?**

SELECT VEHICLE\_TYPE, COUNT (\*) AS NO\_OF\_RIDES\_BOOKED\_FOR\_VEHICLE

FROM FINAL\_BOOKINGS

GROUP BY VEHICLE\_TYPE

ORDER BY NO\_OF\_RIDES\_BOOKED\_FOR\_VEHICLE DESC

LIMIT 1;

1. **Which vehicle type has the highest average driver rating?**

SELECT

VEHICLE\_TYPE,

ROUND(CAST(AVG(DRIVER\_RATING) AS NUMERIC), 3) AS AVERAGE\_RATING

FROM FINAL\_BOOKINGS

WHERE DRIVER\_RATING IS NOT NULL -- Exclude NULL ratings

GROUP BY VEHICLE\_TYPE

ORDER BY AVERAGE\_RATING DESC

LIMIT 1;

### ****Cancellation & Incomplete Ride Analysis****

1. **What are the most common reasons for ride cancellations by customers?**

SELECT REASON\_FOR\_CANCELLING\_BY\_CUSTOMER, COUNT (\*) AS RIDES\_CANCELLED

FROM FINAL\_BOOKINGS

WHERE REASON\_FOR\_CANCELLING\_BY\_CUSTOMER IS NOT NULL

GROUP BY REASON\_FOR\_CANCELLING\_BY\_CUSTOMER

ORDER BY RIDES\_CANCELLED DESC

LIMIT 5;

1. **What is the most common reason for ride cancellations by drivers?**

SELECT REASON\_FOR\_CANCELLING\_BY\_DRIVER, COUNT (\*) AS RIDES\_CANCELLED

FROM FINAL\_BOOKINGS

WHERE REASON\_FOR\_CANCELLING\_BY\_DRIVER IS NOT NULL

GROUP BY REASON\_FOR\_CANCELLING\_BY\_DRIVER

ORDER BY RIDES\_CANCELLED DESC

LIMIT 5;

1. **What is the most common reason for incomplete rides?**

SELECT INCOMPLETE\_RIDES\_REASON, COUNT (\*) AS RIDES\_CANCELLED

FROM FINAL\_BOOKINGS

WHERE INCOMPLETE\_RIDES\_REASON IS NOT NULL

GROUP BY INCOMPLETE\_RIDES\_REASON

ORDER BY RIDES\_CANCELLED DESC

LIMIT 5;

### ****Location & Demand Analysis****

1. **Which pickup location has the highest number of bookings?**

SELECT PICKUP\_LOCATION, COUNT(\*) AS NO\_OF\_TIMES\_PICKUP\_LOCATION\_CHOOSED

FROM FINAL\_BOOKINGS

GROUP BY PICKUP\_LOCATION

ORDER BY NO\_OF\_TIMES\_PICKUP\_LOCATION\_CHOOSED DESC

LIMIT 1;

1. **Which drop location is the most common?**

SELECT PICKUP\_LOCATION, COUNT(\*) AS NO\_OF\_TIMES\_PICKUP\_LOCATION\_CHOOSED

FROM FINAL\_BOOKINGS

GROUP BY PICKUP\_LOCATION

ORDER BY NO\_OF\_TIMES\_PICKUP\_LOCATION\_CHOOSED DESC

LIMIT 1;

### ****Time-Based Analysis****

1. **What are the peak hours for ride bookings?**

SELECT EXTRACT(HOUR FROM TIME) AS HOURS, COUNT(\*) AS NO\_OF\_RIDES\_BOOKED\_AT\_THIS\_TIME

FROM FINAL\_BOOKINGS

GROUP BY HOURS

ORDER BY NO\_OF\_RIDES\_BOOKED\_AT\_THIS\_TIME DESC

LIMIT 3;

1. **How does the number of bookings vary by weekday vs. weekend?**

SELECT EXTRACT(ISODOW FROM DATE) AS WEEKDAY, TO\_CHAR(DATE, 'DAY') AS WEEKDAY\_NAME, COUNT(\*) AS TOTAL\_BOOKINGS

FROM FINAL\_BOOKINGS

GROUP BY WEEKDAY, WEEKDAY\_NAME

ORDER BY WEEKDAY DESC;

**View of the results (SQL)**

--1. What is the total number of bookings?

SELECT \* FROM total\_number\_of\_bookings;

--2. What is the percentage distribution of all ride statuses?

SELECT \* FROM percentage\_distribution\_of\_all\_ride\_statuses;

--3. How many average bookings were made per day?

SELECT \* FROM average\_bookings\_were\_made\_per\_day;

--4. Which day of the week has the highest number of bookings?

SELECT \* FROM day\_of\_the\_week\_with\_the\_highest\_number\_of\_bookings;

--5. Who are the top 5 customers based on the highest number of bookings?

SELECT \* FROM top\_5\_customers\_based\_on\_the\_highest\_number\_of\_bookings;

--6. Who are the top 5 customers based on the highest total amount paid?

SELECT \* FROM top\_5\_customers\_based\_on\_the\_highest\_total\_amount\_paid;

--7. Which vehicle type is booked the most?

SELECT \* FROM vehicle\_type\_booked\_the\_most;

--8. Which vehicle type has the highest average driver rating?

SELECT \* FROM vehicle\_type\_the\_highest\_average\_driver\_rating

--9. What are the most common reasons for ride cancellations by customers?

SELECT \* FROM most\_common\_reasons\_for\_ride\_cancellations\_by\_customers;

--10. What is the most common reason for ride cancellations by drivers?

SELECT \* FROM most\_common\_reason\_for\_ride\_cancellations\_by\_drivers;

--11. What is the most common reason for incomplete rides?

SELECT \* FROM most\_common\_reason\_for\_incomplete\_rides;

--12. Which pickup location has the highest number of bookings?

SELECT \* FROM pickup\_location\_has\_the\_highest\_number\_of\_bookings;

--13. Which drop location is the most common?

SELECT \* FROM drop\_location\_is\_the\_most\_common;

--14. What are the peak hours for ride bookings?

SELECT \* FROM peak\_hours\_for\_ride\_bookings;

--15. How does the number of bookings vary by weekday vs. weekend?

SELECT \* FROM number\_of\_bookings\_vary\_by\_weekday\_vs\_weekend;

**Power BI Answers**

1. **Ride volume over time -** A time series chart showing the number of rides per day.
2. **Booking status breakdown -** A pie chart displaying the proportion of different booking statuses.
3. **Top 5 vehicle types by ride distance -** A stacked bar/column chart ranking vehicle types based on the distance covered.
4. **Revenue distribution of vehicle types -** A stacked bar/column chart displaying the total revenue based on vehicle type.
5. **Top 5 customers by total booking value -** A leader board visual listing customers who have spent the most on bookings.
6. **Ride distances distribution per day -** A histogram or scatter plot showing the distribution of ride distances for different dates.
7. **Canceled rides reasons -** A pie chart that highlights the common reason for ride cancellation.
8. **Average customer ratings by vehicle type -**  A column chart showing the average customers ratings for different vehicle type.
9. **Driver ratings distribution -** A box or cards plot visualizing the spread of driver ratings for different vehicle type.
10. **Customer vs.** **Driver ratings** **-** A box or cards plot comparing customer and driver ratings for each completed ride analyzing correlations.