



Data Analyst Project

By Premshankar Saini

Linkedin:

<mailto:https://www.linkedin.com/in/premshankar-saini-7331102a1/?subject=Linkedin>

Rapido Data Analyst Project

SQL Questions:-

- Q1). Retrieve all successful Ride.
- Q2). Find the average ride distance for each vehicle type.
- Q3). Get the total number of cancelled rides by customers:
- Q4). Ratio of Successful Vs Cancelled Ride:
- Q5). List the top 5 Ride_id who booked the highest number of rides:
- Q6). Get the number of rides which was done by bike and auto :
- Q7). Find the maximum and minimum Ride_Charge by Services:
- Q8). Retrieve all rides where payment was made using GPay , misc_charge >40 and total_Fare is less than 200:
- Q9). Find the average Ride_Charge and Total Rides in July 2024:
- Q10). List all Cancelled rides along with the destination which was cancelled by more than 1 time:

Power BI Questions:

- 1. Ride Volume Over Time
- 2. Ride Status Breakdown
- 3. Top 5 Vehicle Types by Ride Distance
- 4. Total Rides by Payment Method
- 5. Total Charges by Services
- 6. Cancelled Ride Percentage
- 7. Top 5 Destination by Charges
- 8. Top 10 Most Ride Cancelled Date
- 9. Total Ride Charges by Month & Services
- 10. Total Distance by Date

SQL Question & Answers:-

Q1). Retrieve all successful Ride:

```
CREATE VIEW Successful_Ride AS  
  
SELECT  
  
    *  
  
FROM  
  
    RIDE  
  
WHERE  
  
    RIDE_STATUS = 'COMPLETED';
```

Q2). Find the average ride distance for each vehicle type:

```
CREATE VIEW Average_Distance AS  
  
SELECT  
  
    services, AVG(distance) AS Avg_Distance  
  
FROM  
  
    ride  
  
GROUP BY services;
```

Q3). Get the total number of cancelled rides by customers:

```
CREATE VIEW Total_Cancelled_Ride AS  
  
SELECT  
  
    COUNT(*) AS Cancelled_Ride  
  
FROM  
  
    ride  
  
WHERE ride_status = 'cancelled';
```

Q4). Ratio of Successful Vs Cancelled Ride:

```
CREATE VIEW Ride_Status_Ratio AS  
  
SELECT  
  
    SUM(CASE  
  
        WHEN ride_status = 'completed' THEN 1
```

```

        ELSE 0
    END) / SUM(CASE
        WHEN ride_status = 'cancelled' THEN 1
        ELSE 0
    END) AS Ratio
FROM
    ride;

```

Q5). List the top 5 Ride_id who booked the highest number of rides:

```

CREATE VIEW HighestBooked_Ride_Id AS
SELECT
    ride_id, COUNT(*) AS Total_Ride
FROM
    ride
GROUP BY ride_id
ORDER BY Total_Ride DESC LIMIT 5;

```

Q6). Get the number of rides which was done by bike and auto :

```

CREATE VIEW Ride_By_Bike_Auto AS
SELECT
    services, COUNT(*) AS Total_ride
FROM
    ride
WHERE
    services IN ('bike' , 'auto')
GROUP BY services;

```

Q7). Find the maximum and minimum Ride_Charge by Services:

```

CREATE VIEW Max_Min_Charge AS
SELECT
    services,

```

```
MAX(ride_charge) AS Max_Charge,  
MIN(ride_charge) AS Min_Charge  
FROM  
ride  
GROUP BY services;
```

Q8. Retrieve all rides where payment was made using GPay , misc_charge >40 and total_Fare is less then 200:

```
CREATE VIEW GPay_MiscCharge_TotalFare AS  
  
SELECT  
    *  
  
FROM  
    ride  
  
WHERE  
    payment_method = 'GPay'  
    AND misc_charge > 40  
    AND total_fare < 200;
```

Q9). Find the average Ride_Charge and Total Rides in July 2024:

```
CREATE VIEW July2024_Data AS  
  
SELECT  
    AVG(ride_charge) AS Avg_Charge, COUNT(*) AS Total_Rides  
  
FROM  
    ride  
  
WHERE  
    date LIKE ('%07-2024');
```

Q10). List all Cancelled rides along with the destination which was cancelled by more then 1 time:

```
CREATE VIEW Cancelled_Ride_Destination AS

SELECT
    destination, COUNT(destination) AS Total_times
FROM
    ride
WHERE
    ride_status = 'cancelled'
GROUP BY destination
HAVING Total_times > 1;
```

Retrieve All Questions & Answers(SQL)

Q1). Retrieve all successful Ride:

```
SELECT * FROM SUCCESSFUL_RIDE;
```

Q2). Find the average ride distance for each vehicle type:

```
SELECT * FROM AVERAGE_DISTANCE;
```

Q3). Get the total number of cancelled rides by customers:

```
SELECT * FROM TOTAL_CANCELLED_RIDE;
```

Q4). Ratio of Successful Vs Cancelled Ride:

```
SELECT * FROM RIDE_STATUS_RATIO;
```

Q5). List the top 5 Ride_id who booked the highest number of rides:

```
SELECT * FROM HIGHESTBOOKED_RIDE_ID;
```

Q6). Get the number of rides which was done by bike and auto :

```
SELECT * FROM RIDE_BY_BIKE_AUTO;
```

Q7). Find the maximum and minimum Ride_Charge by Services:

```
SELECT * FROM MAX_MIN_CHARGE;
```

Q8. Retrieve all rides where payment was made using GPay , misc_charge >40 and total_Fare is less than 200:

```
SELECT * FROM GPAY_MISCCHARGE_TOTALFARE;
```

Q9). Find the average Ride_Charge and Total Rides in July 2024:

```
SELECT * FROM JULY2024_DATA;
```

Q10). List all Cancelled rides along with the destination which was cancelled by more than 1 time:

```
SELECT * FROM CANCELLED_RIDE_DESTINATION;
```

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. Charges/Cost

- Total Rides by Payment Method
- Total Charges by Services

- Top 5 Destination by Charges

4. Date/Time

- Top 10 Most Ride Cancelled Date
- Total Ride Charges by Month & Services
- Total Distance by Date
- Cancelled Ride Percentage

Answers:

1). Ride Volume Over Time: A line chart showing the total rides over time from July to August 2024.

2). Ride Status Breakdown: A donut chart displaying the proportion of completed vs. cancelled rides.

3). Top 5 Vehicle Types by Ride Distance: Likely a bar chart comparing distances for different vehicle types.

4). Total Rides by Payment Method: A bar chart illustrating the number of rides by each payment method (Paytm, GPay, etc.).

5). Total Charges by Services: A bar chart showing total ride charges across various services like bike, auto, cab.

6). Cancelled Ride Percentage: A KPI card displaying the overall cancelled ride percentage (10.07%).

7). Top 5 Destinations by Charges: A table listing destinations with the highest ride charges.

8). Top 10 Most Ride Cancelled Dates: A table showing dates with the highest number of cancelled rides.

9). Total Ride Charges by Month & Services: A stacked bar chart displaying ride charges split by month and service type.

10). Total Distance by Date: A line chart depicting total distance covered over time in July and August 2024.