

SQL Relationship dataset Analysis

GitHub : [pradeep00000](#)

```
1 • create schema game;
2 • select * from game.mta_daily_ridership;
3
4 #1. Fetch the Total Subway Ridership on a Specific Date
5 • select `Subways: Total Estimated Ridership`
6   from game.mta_daily_ridership
7  where Date = '2020-03-01';
8
9 #2. List All Unique Dates in the Dataset
10 • select distinct Date
11    from game.mta_daily_ridership;
12
13 #3. Find the Average Daily Bus Ridership
14 • select Date , avg(`Buses: Total Estimated Ridership`)
15    from game.mta_daily_ridership
16   group by Date;
```



```
18 #4. Show the Maximum Subway Ridership Recorded
19 • select max(`Subways: Total Estimated Ridership`) as Relationship
20 from game.mta_daily_ridership;
21
22 #5. Count the Total Number of Days in the Dataset
23 • select distinct count(Date)
24 from game.mta_daily_ridership;
25
26 #6. Calculate the Total Ridership Across All
27 # Transit Modes on a Specific Date
28 • select Date, `Subways: Total Estimated Ridership`
29 + `Buses: Total Estimated Ridership`
30 + `LIRR: Total Estimated Ridership`
31 + `Metro-North: Total Estimated Ridership`
32 + `Staten Island Railway: Total Estimated Ridership`
33 as Total_Ridership from game.mta_daily_ridership
34 where Date = '2020-03-01';
```

```
36 #7. Find the Day with the Lowest Subway Ridership
37 • select Date , `Subways: Total Estimated Ridership`
38     as Min_Relationship_Time
39 from game.mta_daily_ridership
40 order by Min_Relationship_Time asc
41 limit 1;
42
43 #8. Rank Days by Total Traffic on Bridges and Tunnels
44 • select Date, `Bridges and Tunnels: Total Traffic` ,
45     rank() over( order by `Bridges and Tunnels: Total Traffic` desc)
46 as 'Rank' from game.mta_daily_ridership;
47
```

```
48 #9. Identify Dates Where Subway Ridership Exceeded Bus Ridership
49 • select Date, `Buses: Total Estimated Ridership`,
50    `Subways: Total Estimated Ridership`
51 from game.mta_daily_ridership
52 where `Buses: Total Estimated Ridership`
53    < `Subways: Total Estimated Ridership`;
54
55 #10. Calculate the Percentage Change in Subway Ridership from One Day to the Next
56 • select Date, `Subways: Total Estimated Ridership`,
57    lag(`Subways: Total Estimated Ridership`) over (order by Date) as Previous_Day_Ridership,
58    ((`Subways: Total Estimated Ridership` - lag(`Subways: Total Estimated Ridership`)
59    over (order by Date)) / lag(`Subways: Total Estimated Ridership`)
60    over (order by Date)) * 100 as Percentage_Change
61 from game.mta_daily_ridership;
62
```



```
63 #11.Find Days Where Total Traffic on Bridges and Tunnels Exceeded the Average by 20%
64 • with AvgTraffic as (
65     select avg(`Bridges and Tunnels: Total Traffic`) as Avg_Traffic
66     from game.mta_daily_ridership
67 )
68 select Date, `Bridges and Tunnels: Total Traffic`
69 from game.mta_daily_ridership, AvgTraffic
70 where `Bridges and Tunnels: Total Traffic` > Avg_Traffic * 1.2;
71
72 #12. Find the Top 5 Days with the Highest Total Ridership
73 • select Date,
74     `Subways: Total Estimated Ridership` + `Buses: Total Estimated Ridership` +
75     `LIRR: Total Estimated Ridership` + `Metro-North: Total Estimated Ridership` +
76     `Staten Island Railway: Total Estimated Ridership` as Total_Ridership
77 from game.mta_daily_ridership
78 order by Total_Ridership desc
79 limit 5;
```

```
81  #13. Analyze the Impact of Pandemic Using Pre-Pandemic Comparisons
82 • select Date,
83     `Subways: % of Comparable Pre-Pandemic Day`,
84     `Buses: % of Comparable Pre-Pandemic Day`,
85     `LIRR: % of Comparable Pre-Pandemic Day`,
86     `Metro-North: % of Comparable Pre-Pandemic Day`
87 from game.mta_daily_ridership
88 where `Subways: % of Comparable Pre-Pandemic Day` < 50;
89
```

```
90 #14. Identify the Mode of Transit with the Largest Drop During the Pandemic
91 • select Date,
92     `Subways: Total Estimated Ridership`,
93     `Buses: Total Estimated Ridership`,
94     `LIRR: Total Estimated Ridership`,
95     `Metro-North: Total Estimated Ridership`
96 from game.mta_daily_ridership
97 where Date between '2020-03-01' and '2020-12-31'
98 order by `Subways: % of Comparable Pre-Pandemic Day` asc
99 limit 1;
```



```
101 #15. Highlight Anomalous Days with Extremely Low Ridership
102 • select- Date, `Subways: Total Estimated Ridership`, `Buses: Total Estimated Ridership`
103 from game.mta_daily_ridership
104 where `Subways: Total Estimated Ridership` < 1000000;
105
```