

DATA Analyst project

By sagar raj yadav

**as a data analyst**

**step 1:data cleaning in excel**

i)go to **Data** ->click **Remove duplicates** (Ribbon ke andar Data wala section hoga)

**step 2:setup sql**

ii)open sql workbench-> CREATE DATABASE database\_name -> USE database\_name

iii)right click on Table->select Table Data Import wizard ->load csv file and wait kro import hone ka.

**step 3: write sql query**

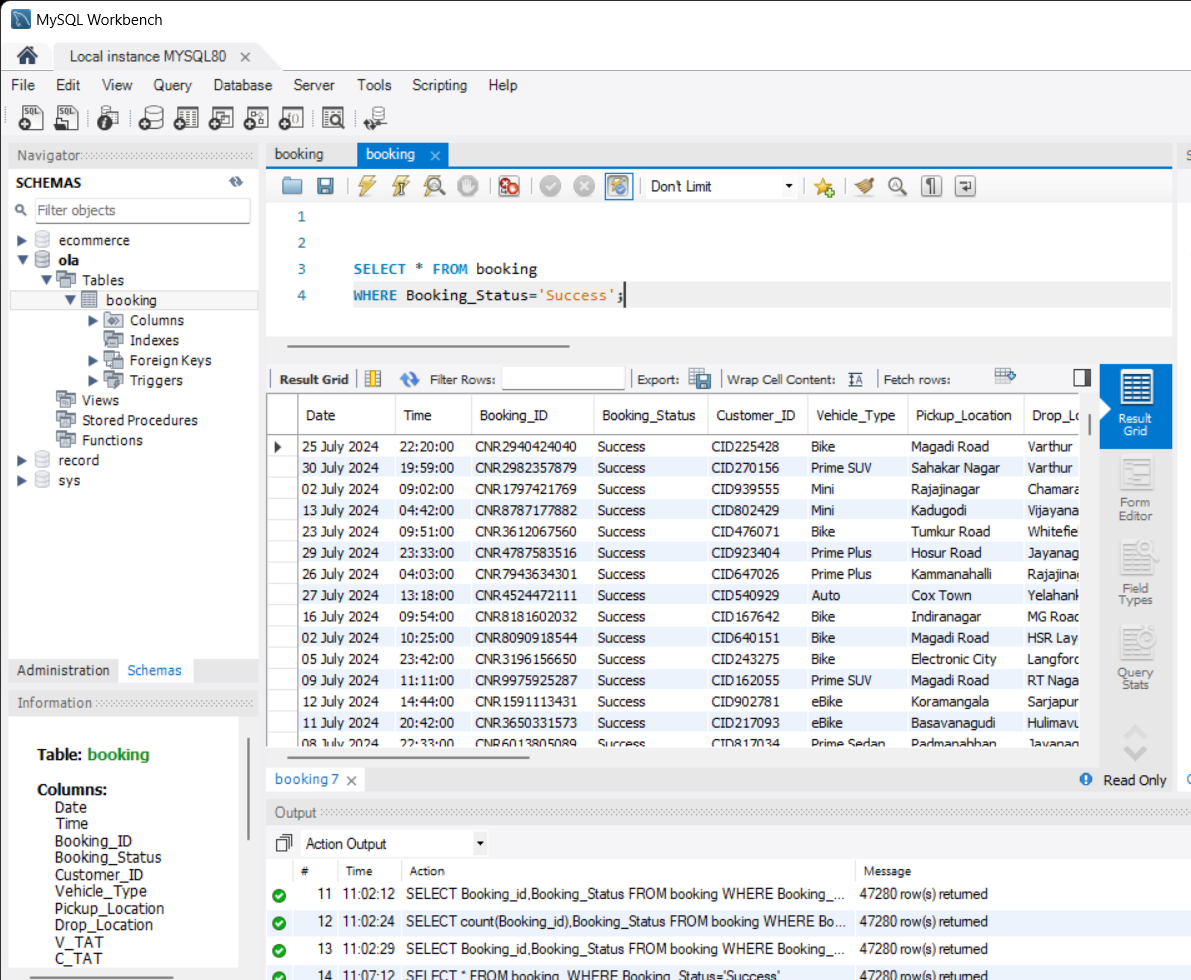
**SQL Questions:**

1. Retrieve all successful bookings

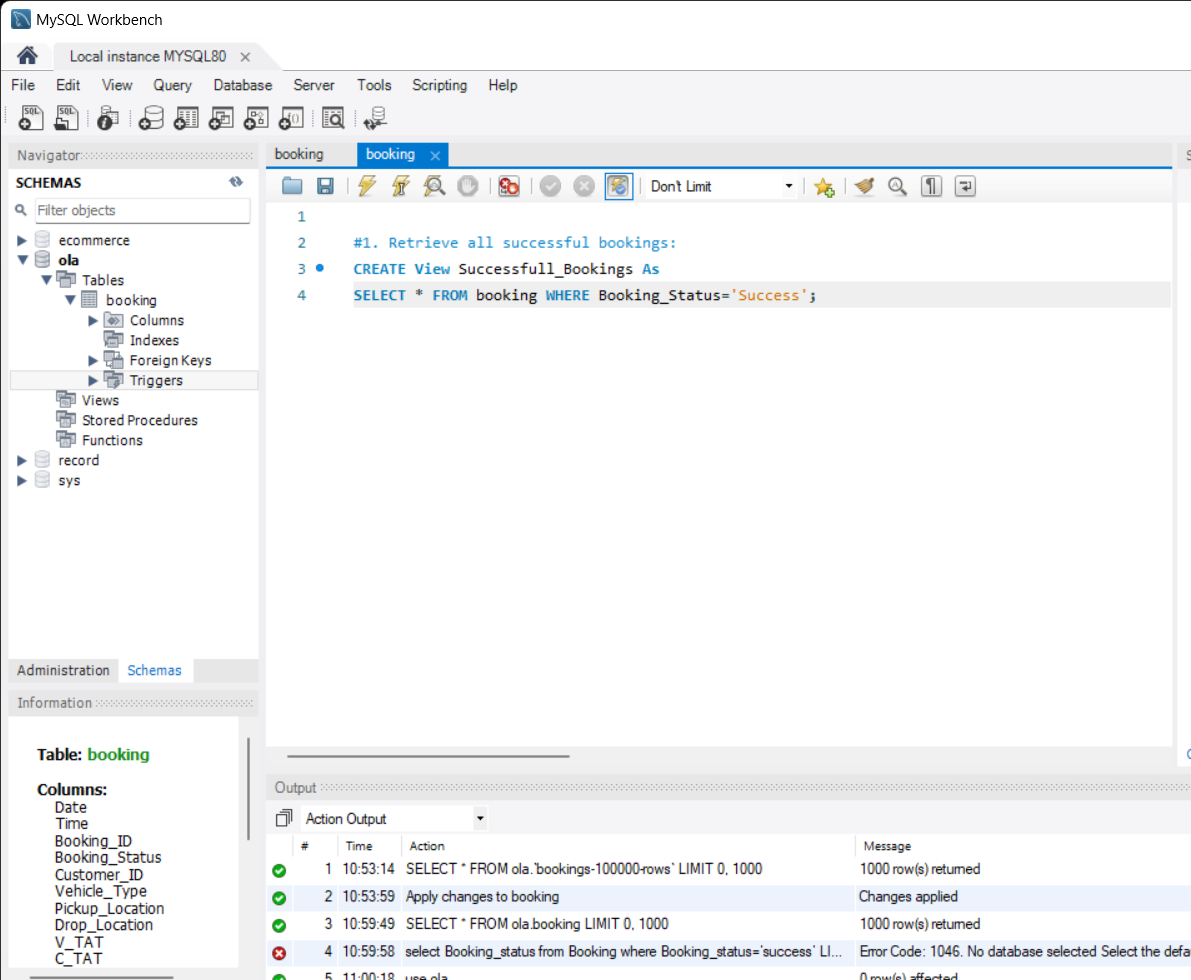
i) **fetch only booking\_id**



ii) **fetch all column**



**iii)create view**

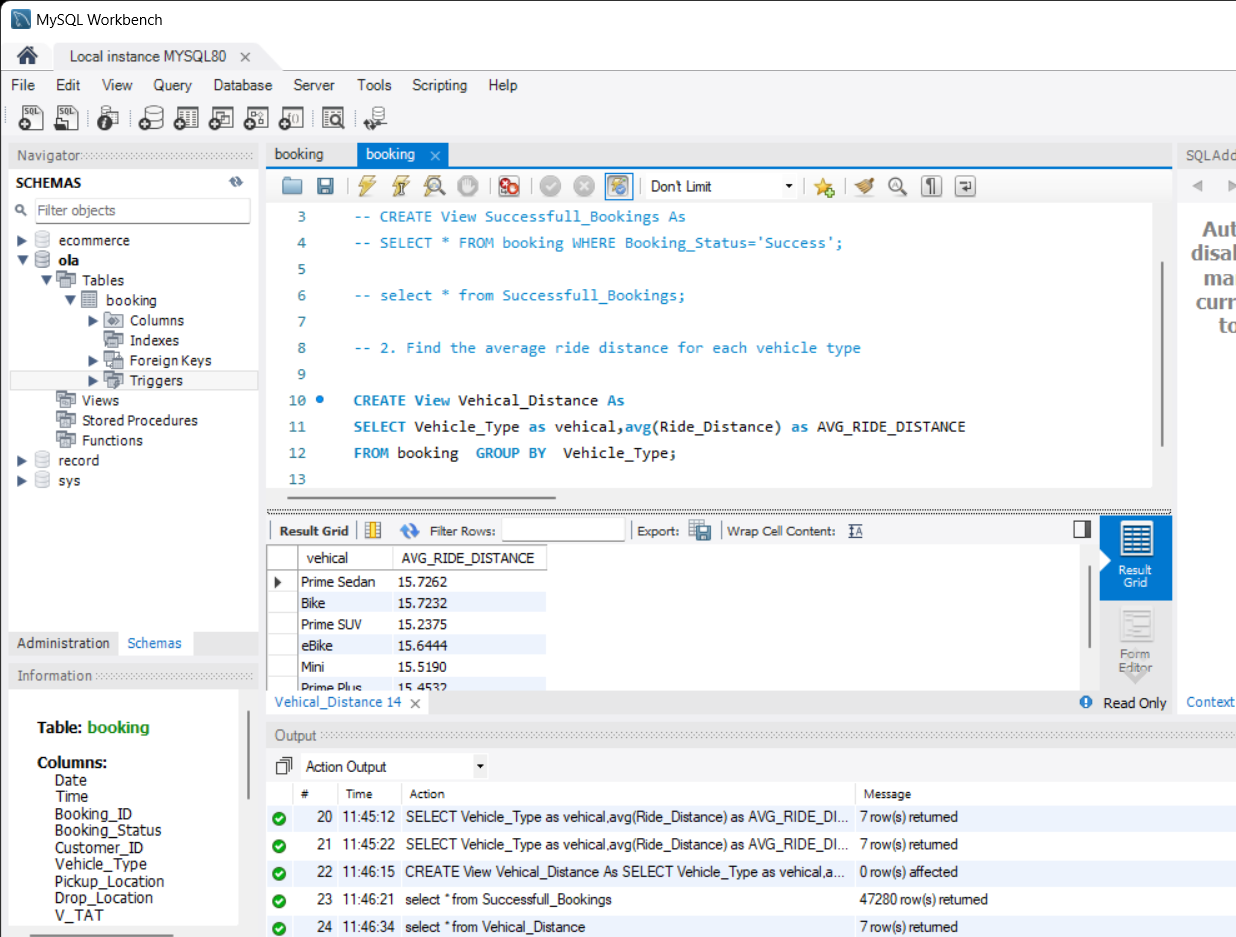


note:-> **select \* from Successfull\_Bookings; -> fetch data from view**

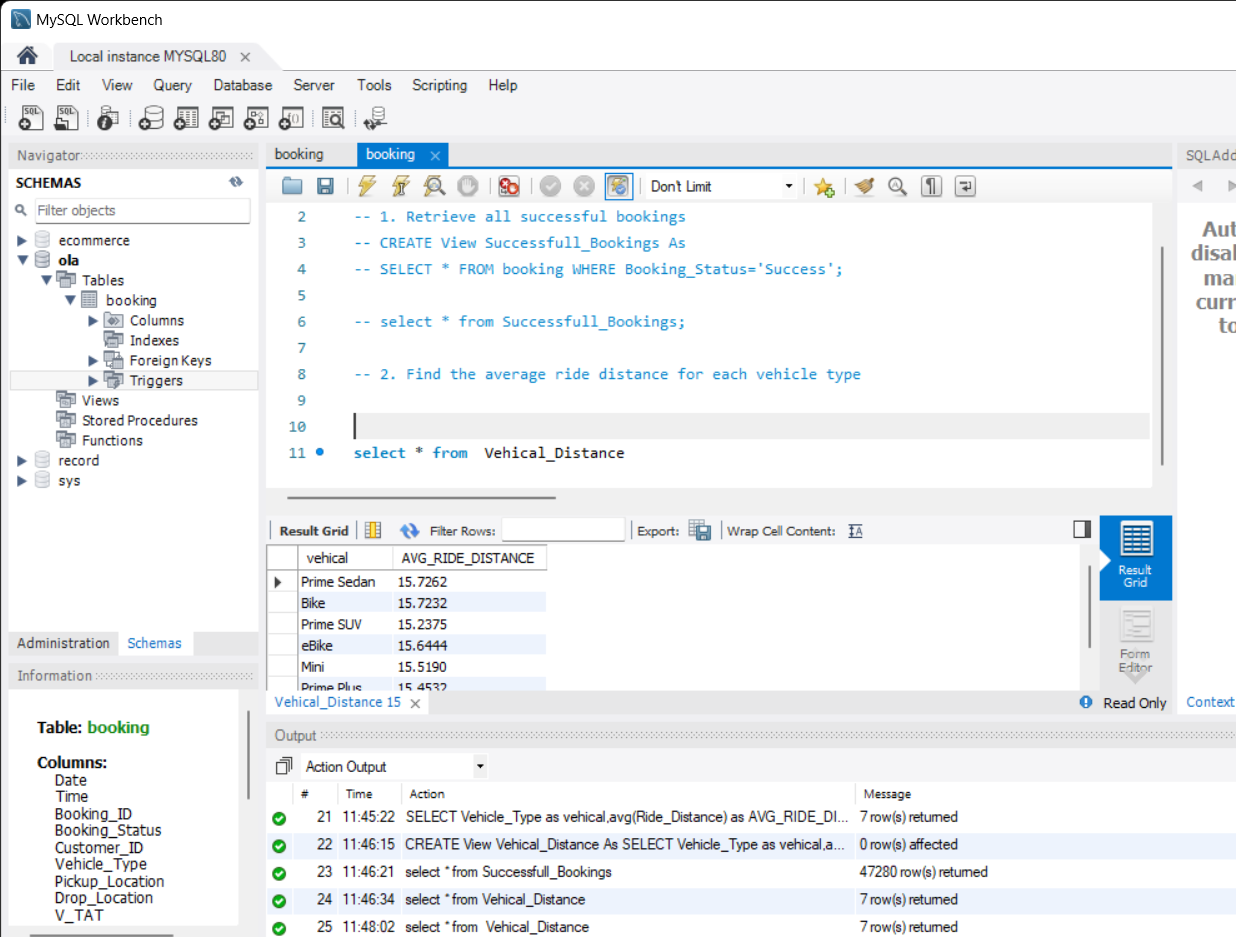
=> View banane ka maksad hai ek query ko reusable bana dena. Jab hum ek **view** banate hain, toh hum ek predefined SQL query ko ek naam de dete hain, jise baad mein hum bar-bar likhne ki bajaye us naam ka use karke data retrieve kar sakte hain.

1. Find the average ride distance for each vehicle type

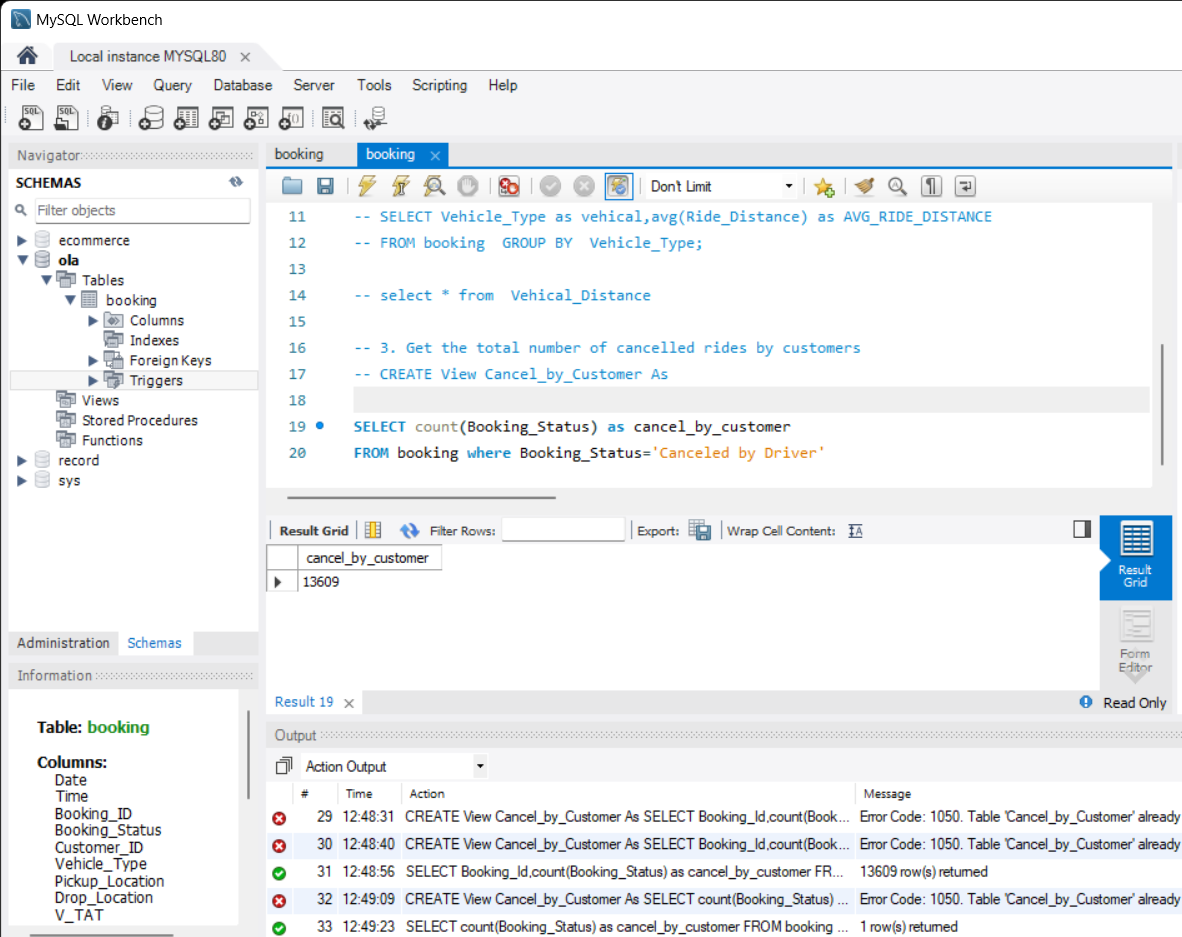
i)



**ii)fetch data from view**



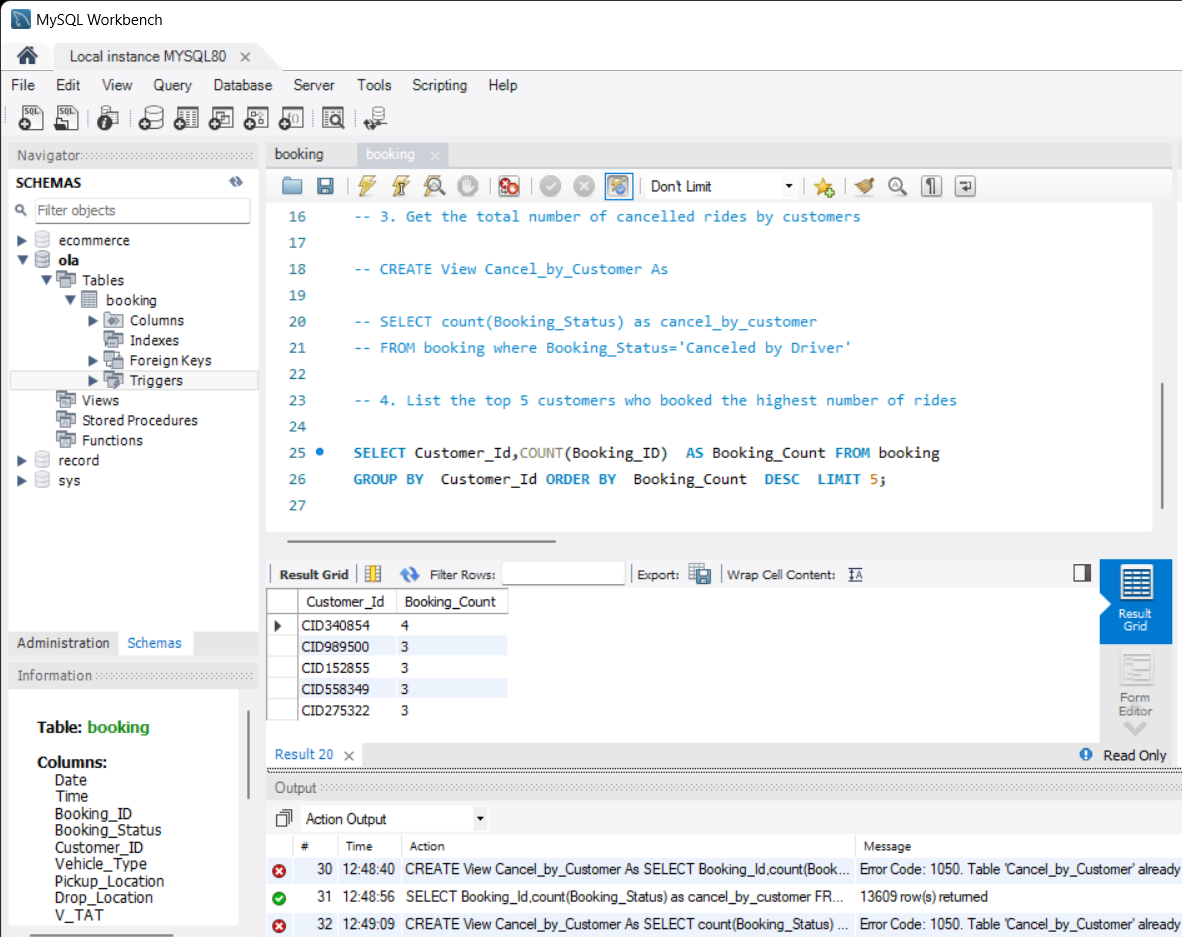
3. Get the total number of cancelled rides by customers:



**similarly create a view,**

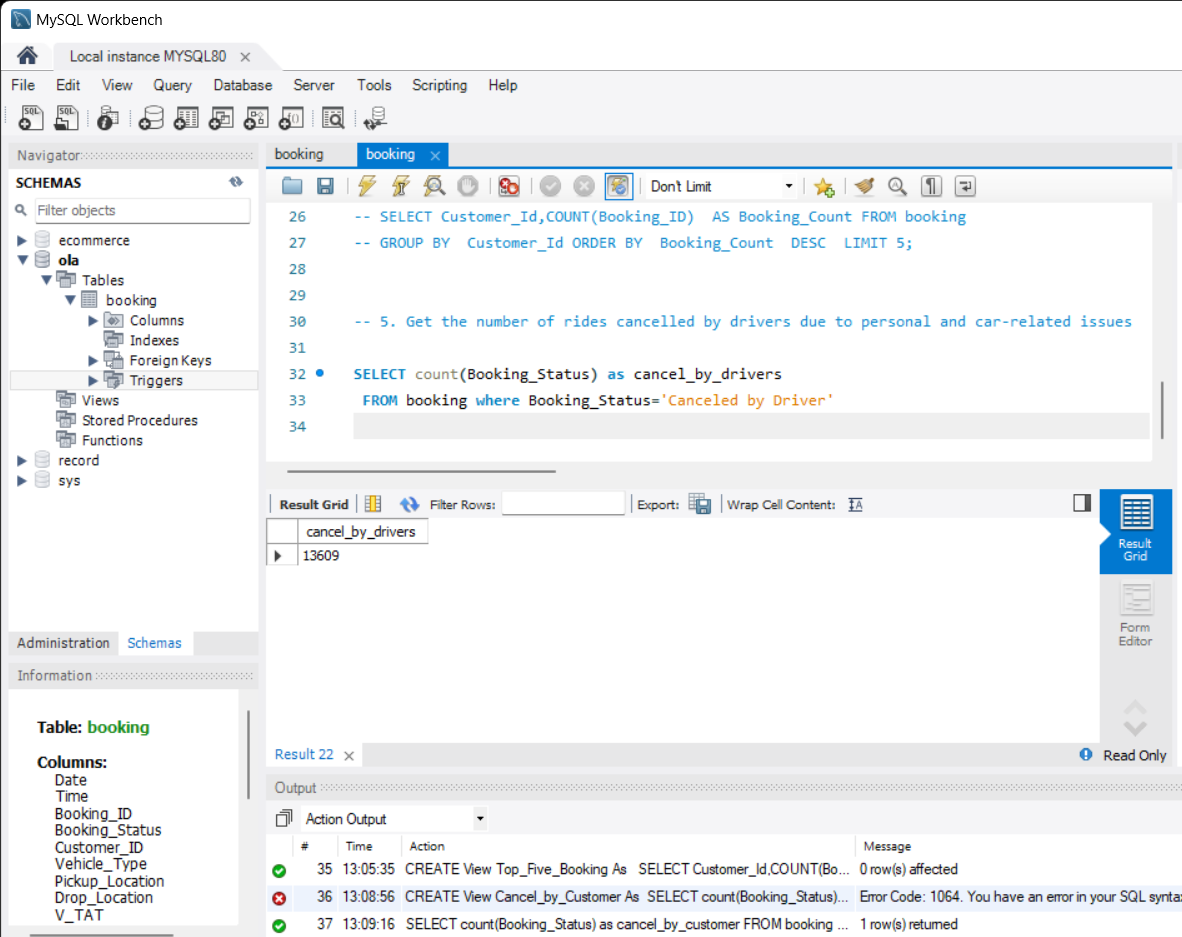


4. List the top 5 customers who booked the highest number of rides:



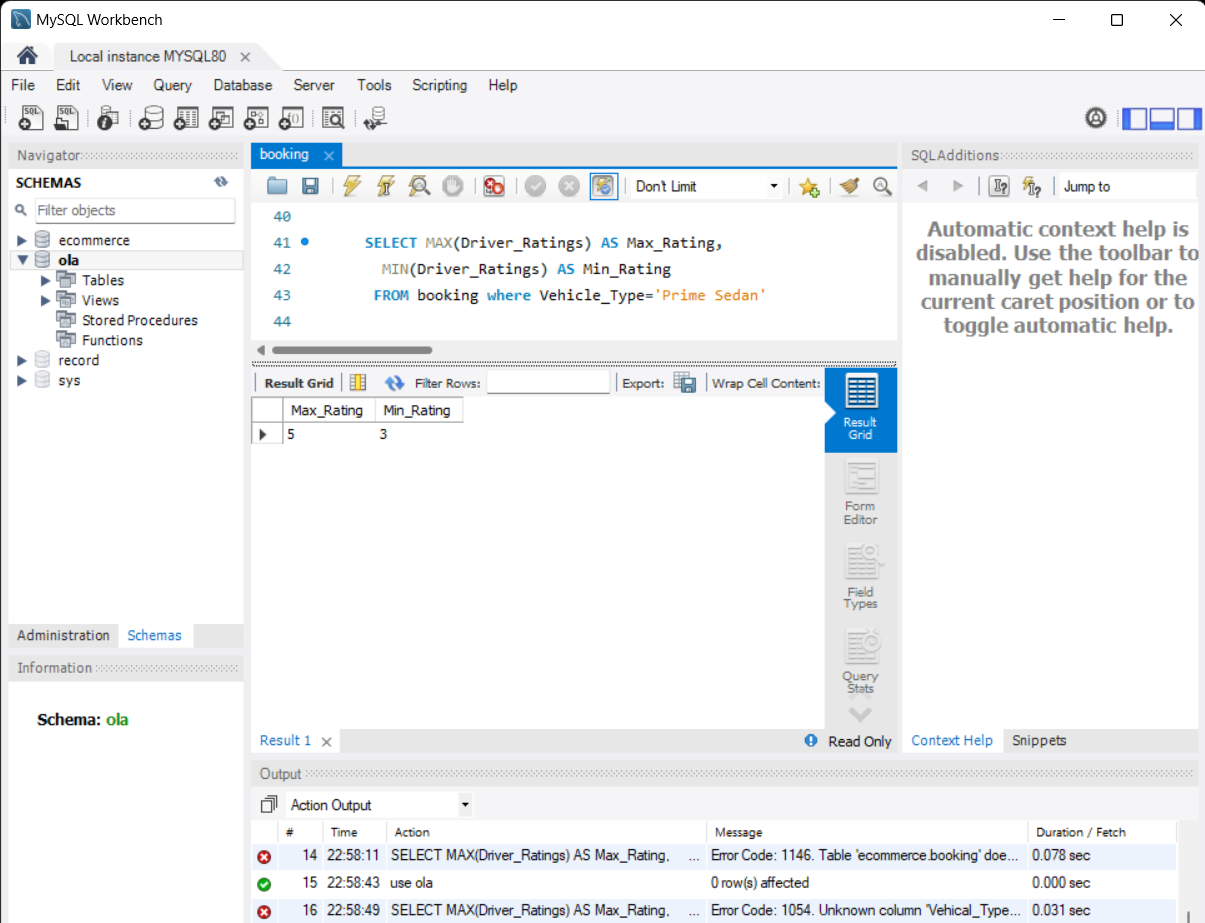
Note: **CREATE View Top\_Five\_Booking As**

5. Get the number of rides cancelled by drivers due to personal and car-related issues:



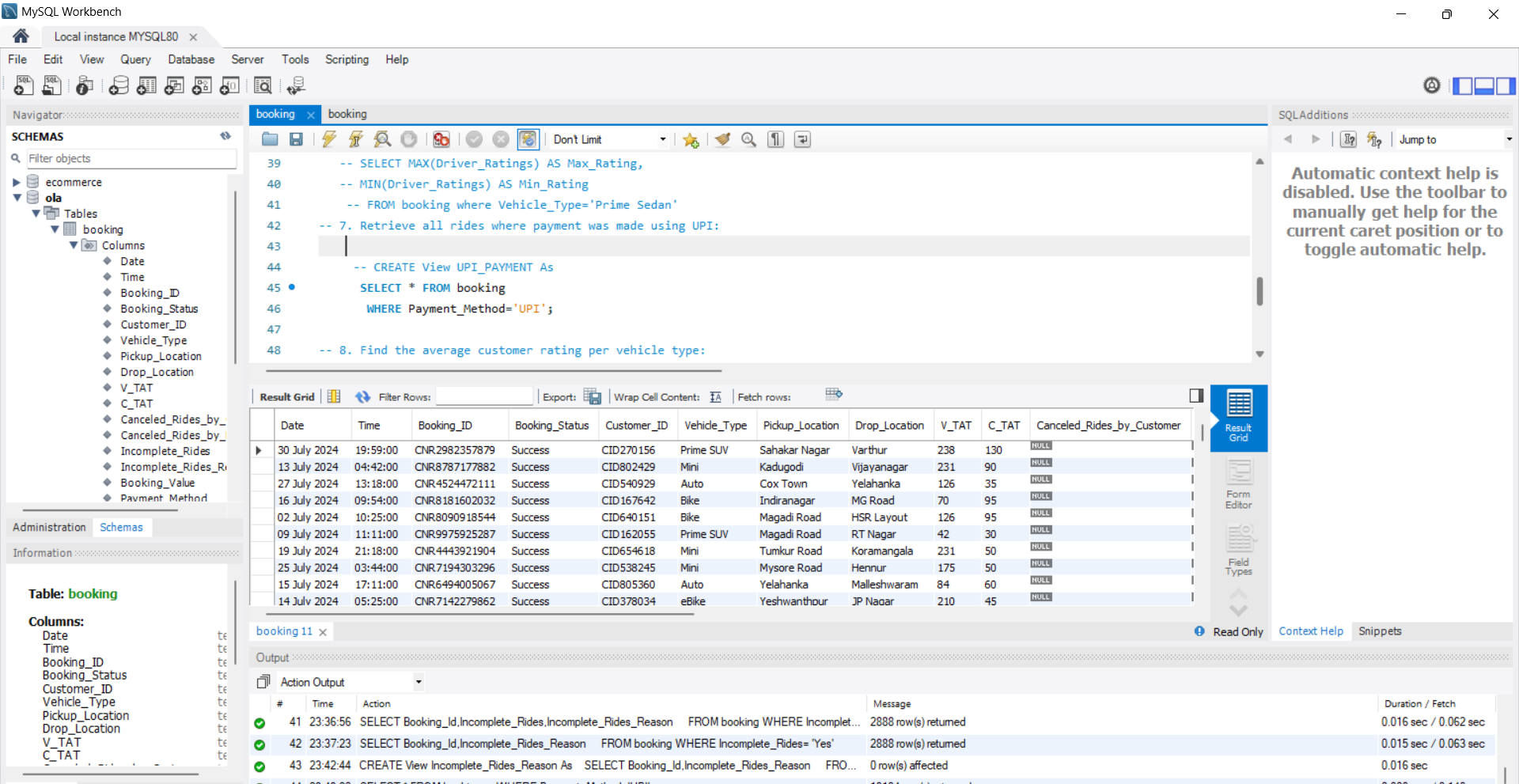
Note: **CREATE View cancel\_by\_drivers As**

6. Find the maximum and minimum driver ratings for Prime Sedan bookings:

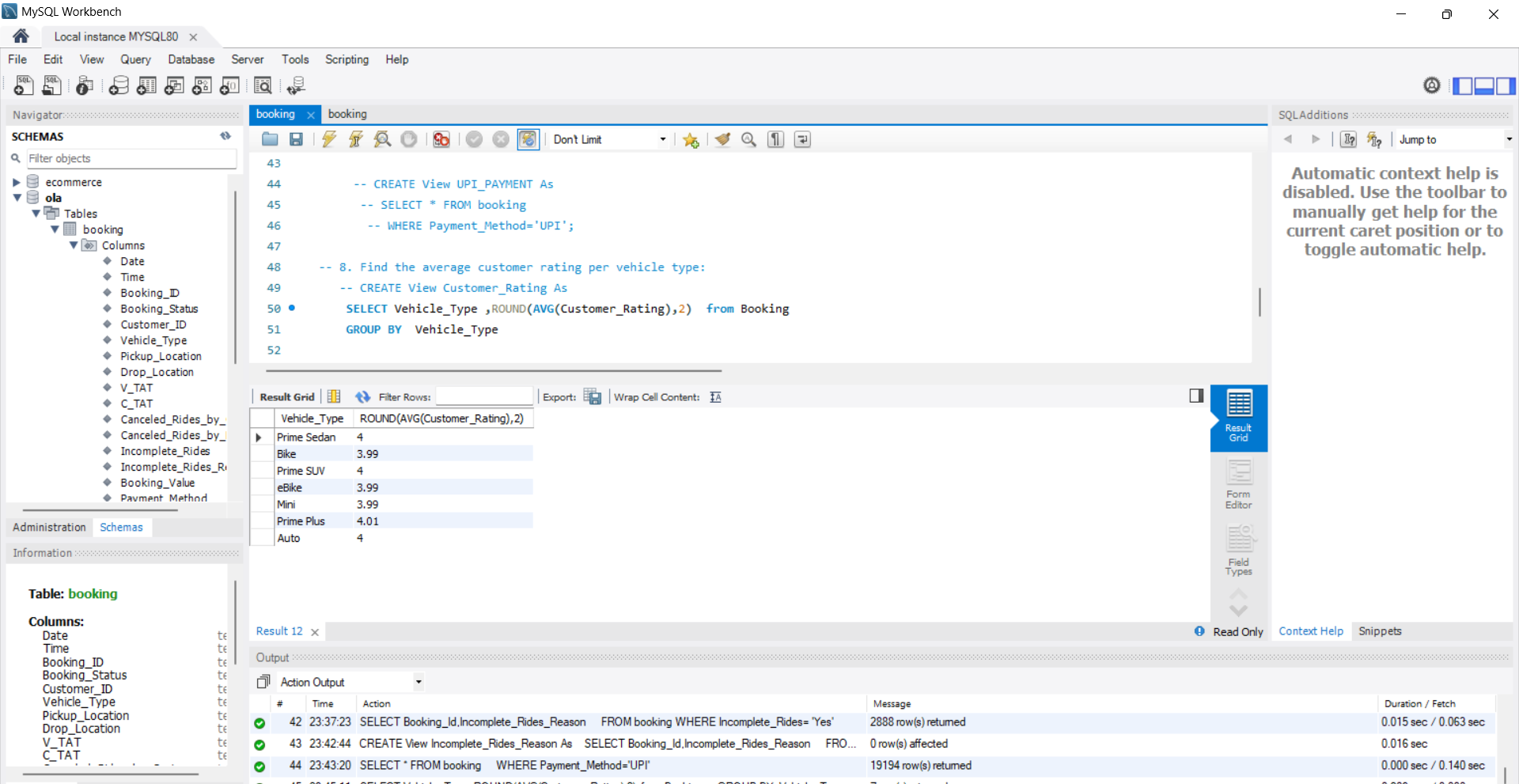


note **: CREATE View Driver\_Rating As**

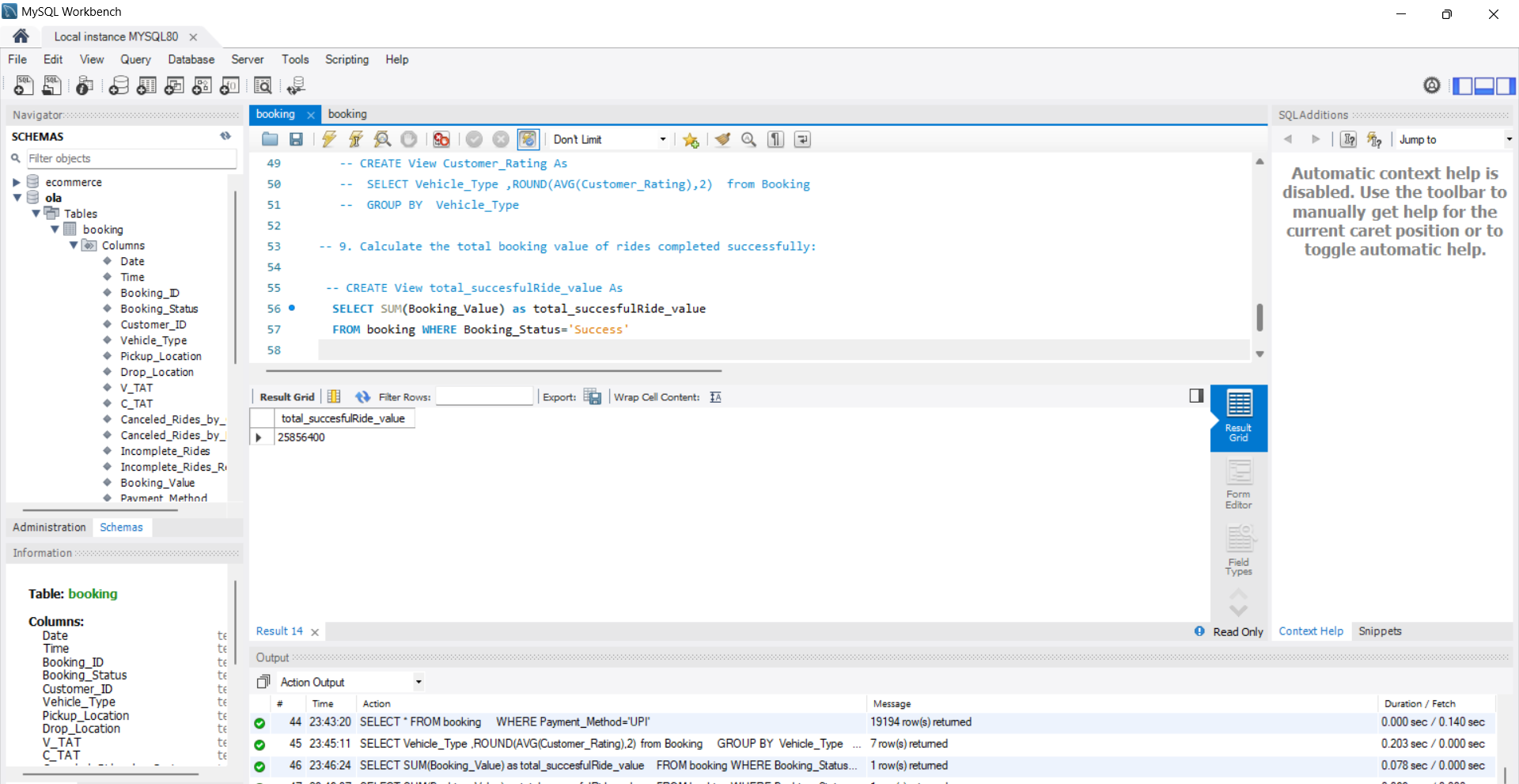
7. Retrieve all rides where payment was made using UPI:



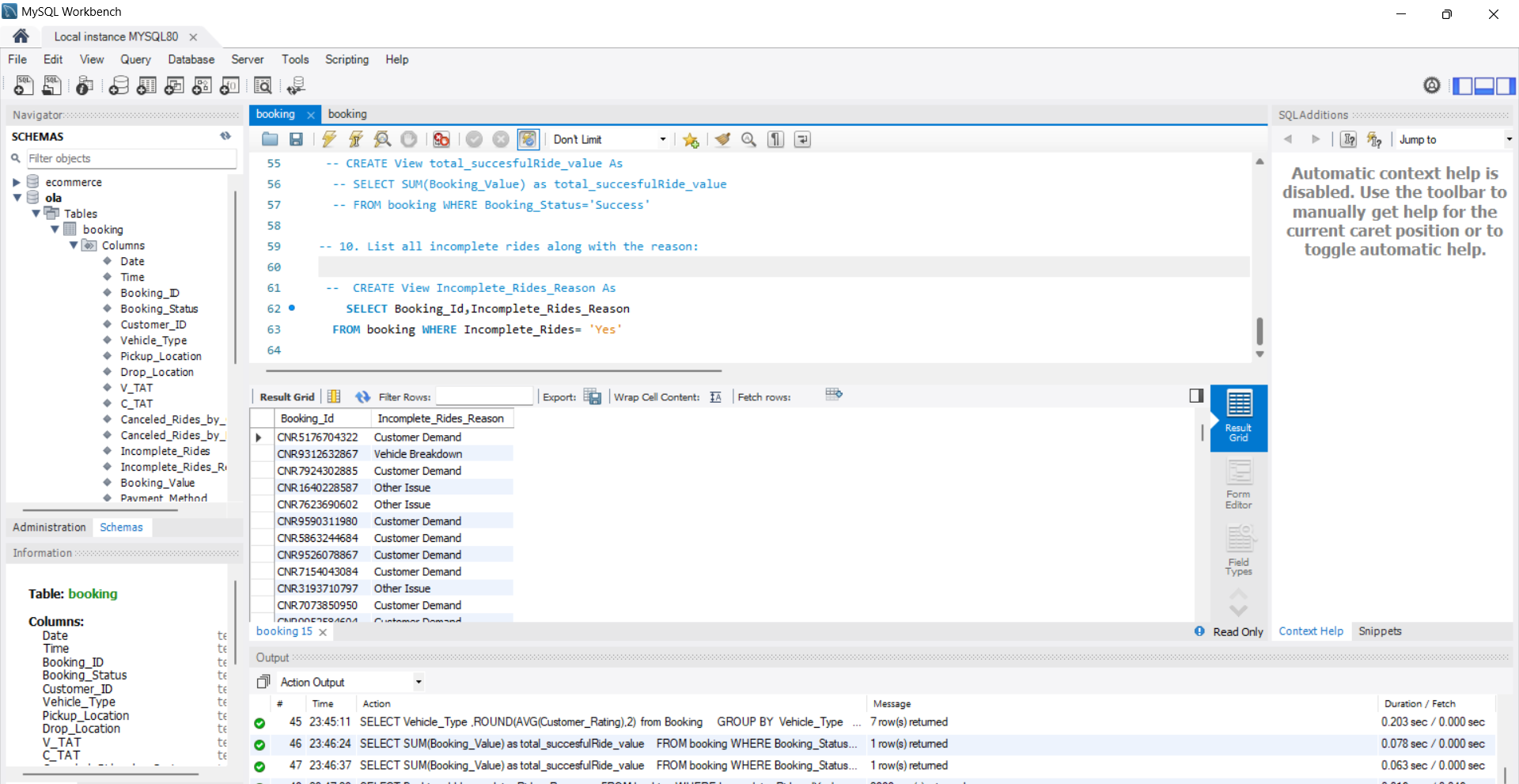
8. Find the average customer rating per vehicle type:



9. Calculate the total booking value of rides completed successfully:



10. List all incomplete rides along with the reason:



**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

**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