BIG DATA
PROJECT 2021

GROUP k

Date: MAY 28, 2021

Our Team

Sukhman

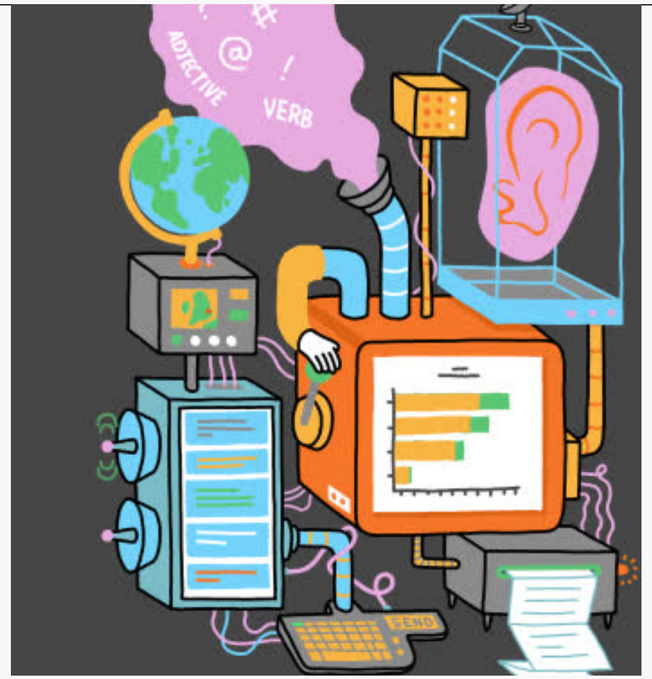
Bhumanyu Singh

Varun Nanda

Himanshu Dhahana

Muhammad Arsalan Alam

Reports Highlights



- Database Design Details
- ER Diagrams
- SQL Commands
- Evidence And Question
- SQL QUERIES
- SCREENSHOTS

TYPES OF DATABASE

There are various types of data

1. Structural Database
2. Free-Form Database
3. Operational Database
4. Analytical Database
5. Rational Database

For this project, we have used Relational Database.

1. DEFINITION: A database structured to recognize relations between stored items of information.

Types of relations in a Relational Database

There are three types of relation as follows

1.One-to-One

A row in a table can have only one matching row in the second table, and vice versa.

2.One-to-Many (or Many-to-One)

A row in one table can have many matching rows in other, but a row in second table can have only one matching row in the first one.

3.Many-to-Many

In a many-to-many relationship, a row in first table can have many matching rows in second table, and vice versa.

Details of Database Design

	Table 1 Cars	Table 2 Seller	Table 3 Dealers
We have created three tables			
Each Table has attributes	Sales id ,name,year,km_driven,selling_price,fuel,mileage,engine,max_power, torque and seats	Seller_id,Seller_type ,and owner	Seller_id and dealer
Relationships between tables		Seller table has many to many relation with cars table	Cars table has many to one relation with dealers table
Most used commands	CREATE TABLE ,SELECT, SELECT,UPDATE,DELETE,INSERT INTO,CREATE DATABASE,ALTER DATABASE,DROP TABLE.		

OUR FEATURES

FEATURE 1

We have created a primary key - Sales_id in the table Cars and this act as foreign key in rest of the two tables

FEATURE 2

The database has been cleaned before working and contains no null values

FEATURE 3

We have generated total_revenue using SUM command

FEATURE 4

We have calculated minimum selling price using SELECT MIN() command

FEATURE 5

We have calculated minimum mileage using SELECT MIN() command

SQL Commands

SELECT- extracts data from a database
UPDATE- updates data in a database
DELETE- deletes data from a database
INSERT INTO- inserts new data into a database
CREATE DATABASE- creates a new database
CREATE TABLE- creates a new table
ALTER TABLE- modifies a table
DROP TABLE- deletes a table
AVG- returns the average value for a numeric column.
As- rename a column or table using an alias.
WHERE- filter the result set to include only rows
AND is an operator that combines two conditions.
COUNT- counts the number of rows where the column is not NULL.

GROUP BY- used with aggregate functions.

SUM- is a function that takes the name of a column as an argument and returns the sum of all the values in that column.

Inner JOIN- will combine rows from different tables if the join condition is true.

ORDER BY- sort the result set by a particular column

LIMIT- specify the maximum number of rows

MIN- returns the smallest value in that column.

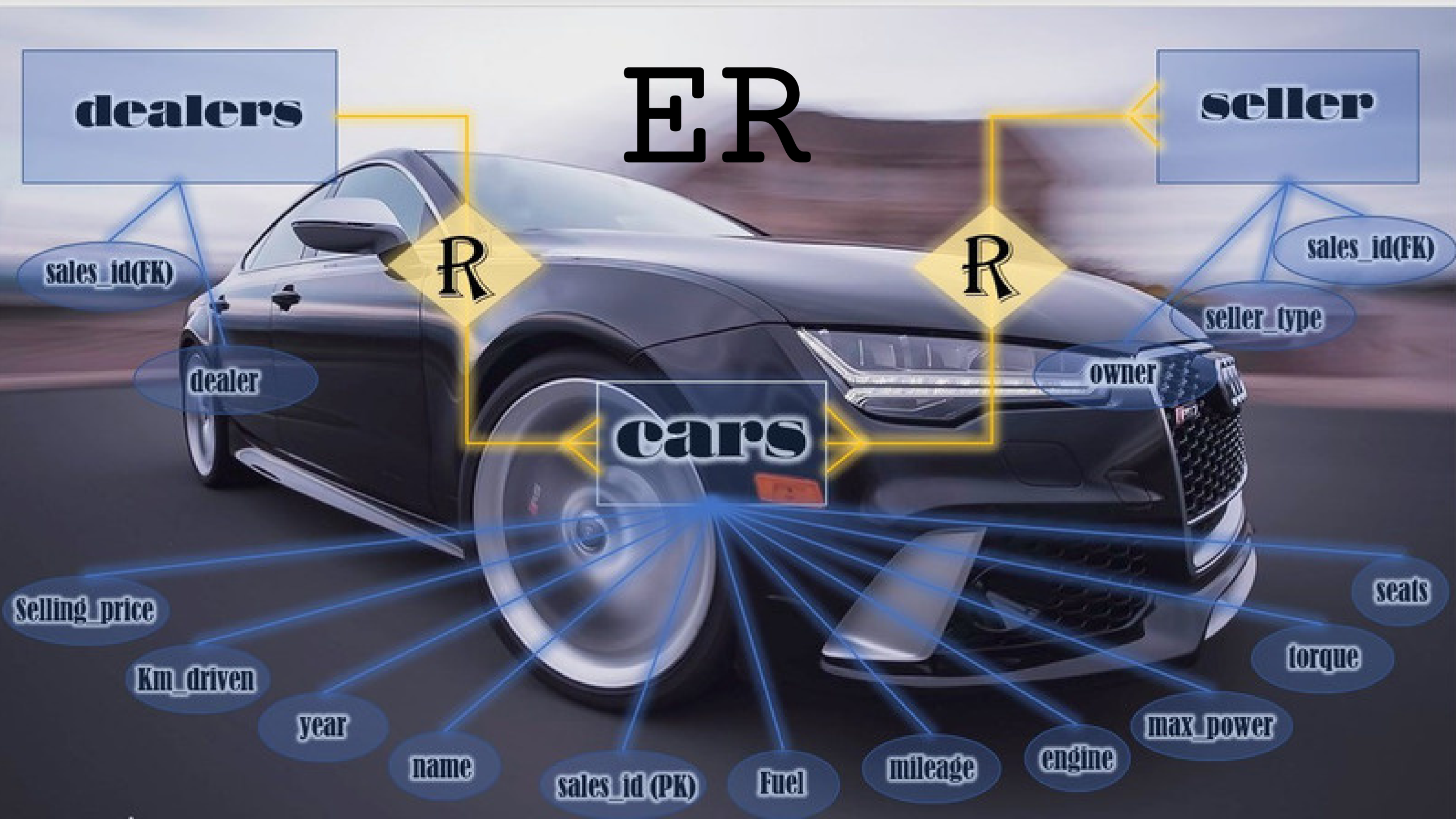
MAX- returns the largest value in that column.

OR- filters the result set

ER DIAGRAM explanation

- Entities are shown with rectangles and attributes with ovals shape .
- The relational line joining cars table to dealers table has three joining lines at cars table and one at dealers table, this shows the MANY to ONE relation .
- The relational line joining cars table to seller table has three joining lines at both tables, this shows the MANY to MANY relation .
- The R in diamond shape shows that there is a relation between tables.

ER DIAGRAM →


ER
dealers
seller
R
R
cars
sales_id(FK)
dealer
sales_id(FK)
seller_type
owner
Selling_price
Km_driven
year
name
sales_id (PK)
Fuel
mileage
engine
max_power
torque
seats

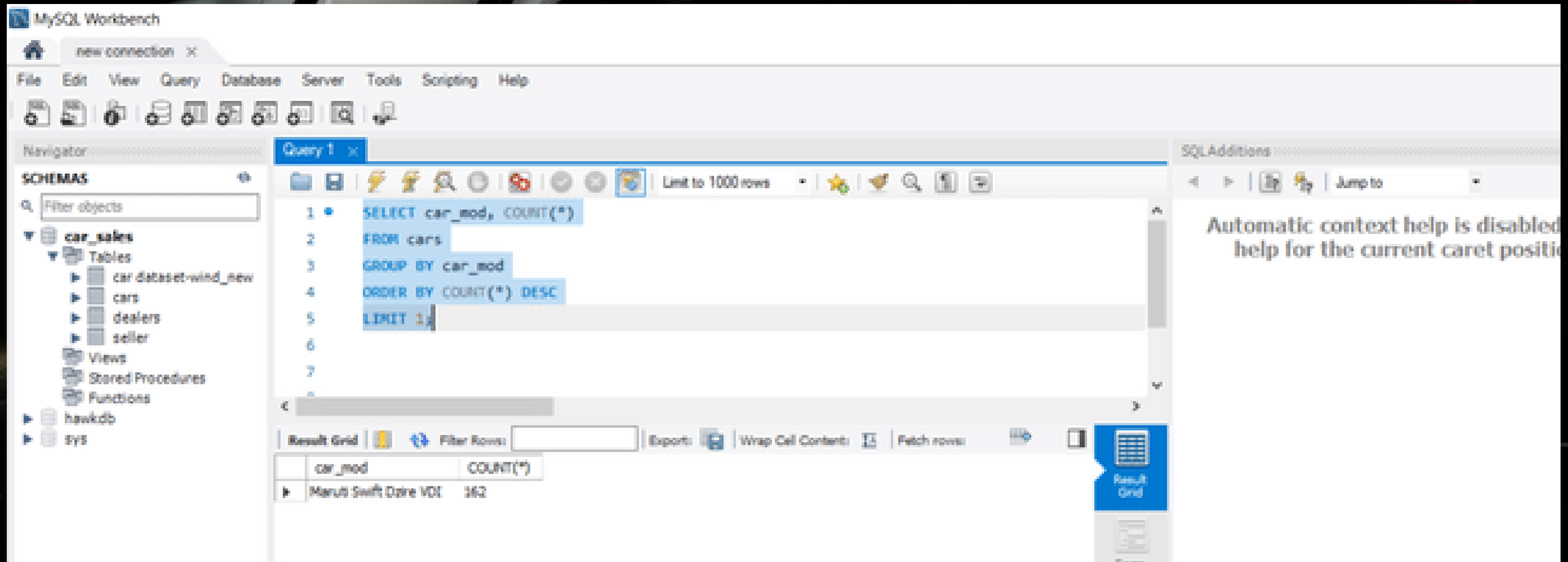
Evidence And QUESTION

NEXT →

Find out which car model sold the most ?

1. We have given the command SELECT car_mod,COUNT(*) this command will count total cars of each type from cars table,
2. Then we have given the command ORDER BY COUNT(*) DESC which will arrange it in descending order.
3. We have added LIMIT 1 which will give us the value on the top

CONCLUSION - MARUTI SUZUKI SWIFT DESIRE VDI WAS SOLD THE MOST (162).

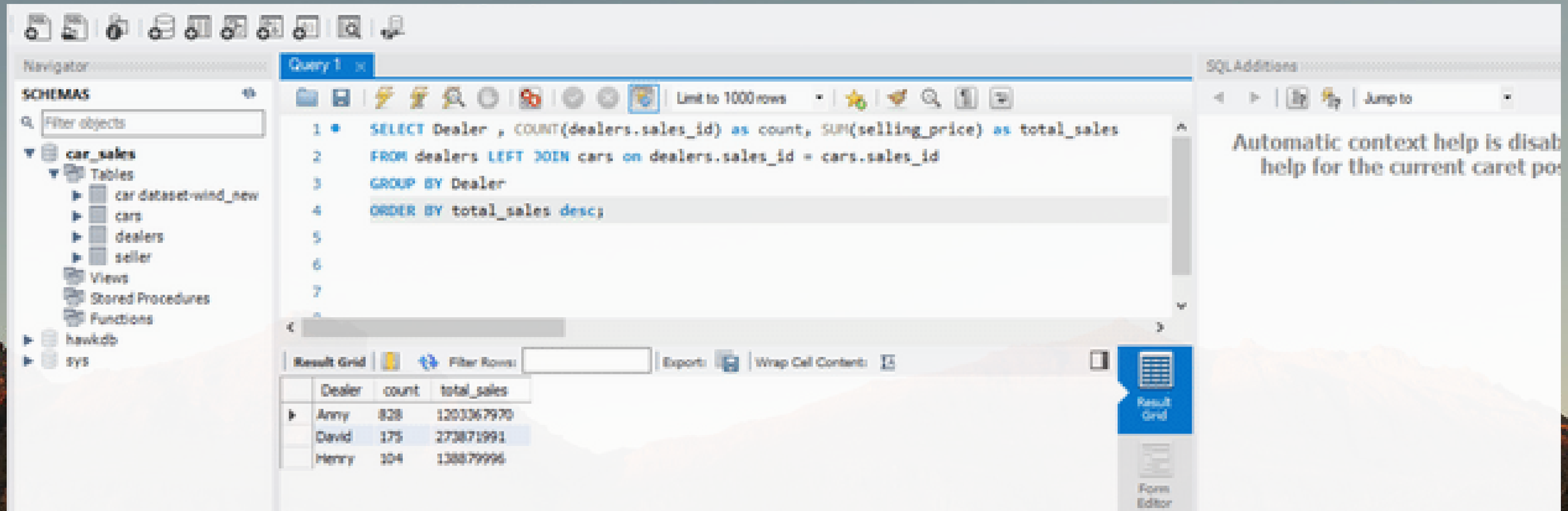


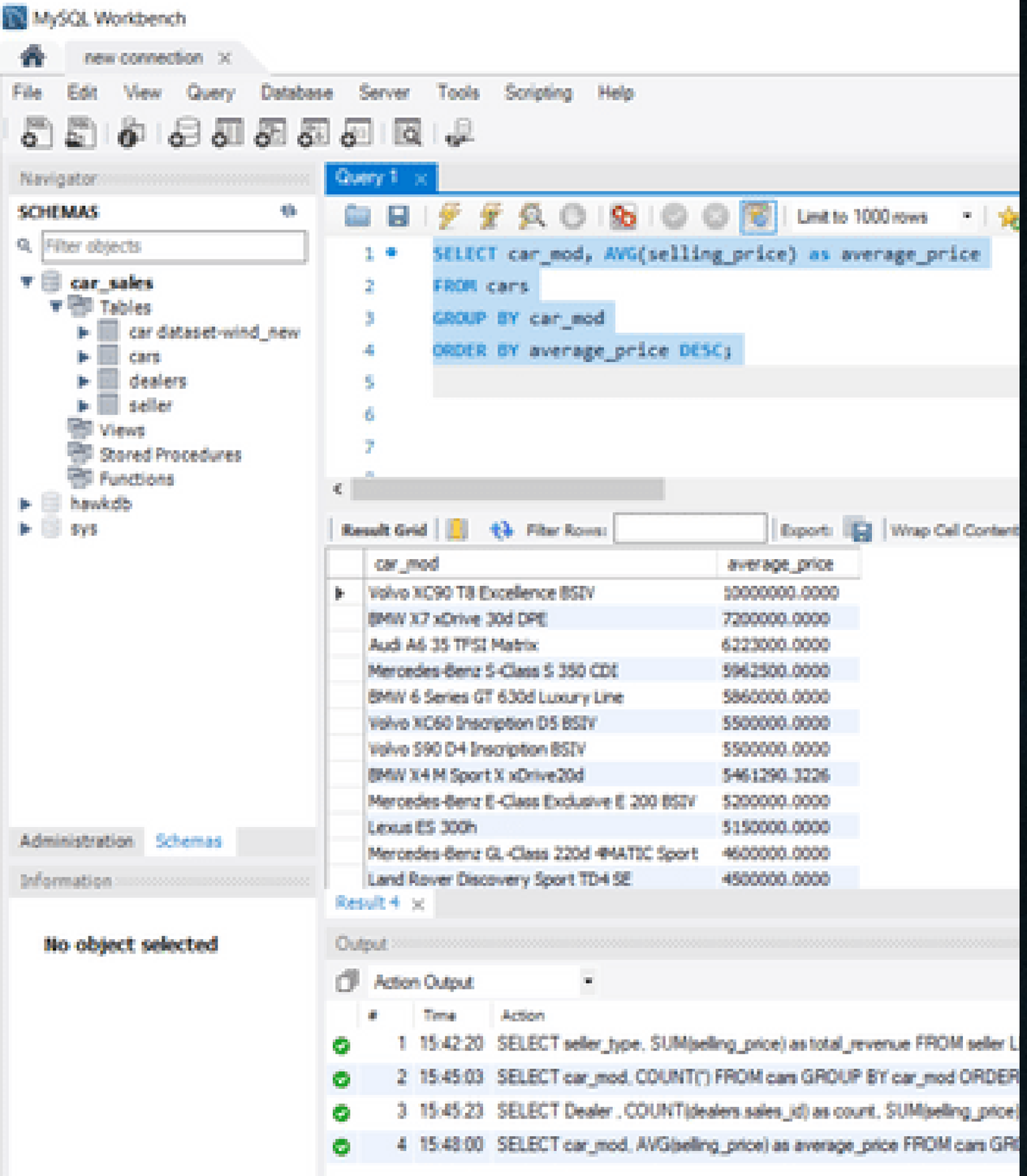
Which dealers' sold more cars, and what is his total sales?

1. We have used the command SELECT Dealer , COUNT (dealers.sales_id) as count,SUM(selling_price) as total_sales this command will count the number of dealers ,will sum up the selling price from dealers table .
2. LEFT JOIN will join the tables cars and dealers .

3. ORDER BY total_sales will tell us the total amount of sales by each dealer

CONCLUSION - The MOST SALES ARE DONE BY ANNY. SHE DID 828 DEALS FOR 1203367970.00





What is the average price for each car model?

1. We have given the command SELECT car_mod, AVG(selling_price) as average_price from cars table,this will calculate the average selling price .

2. We have used the command ORDER BY average_price DESC; this will arrange the average selling price of each mode in descending order.

CONCLUSION - CAN BE CALCULATED AS SHOWN IN THE SCREENSHOT.

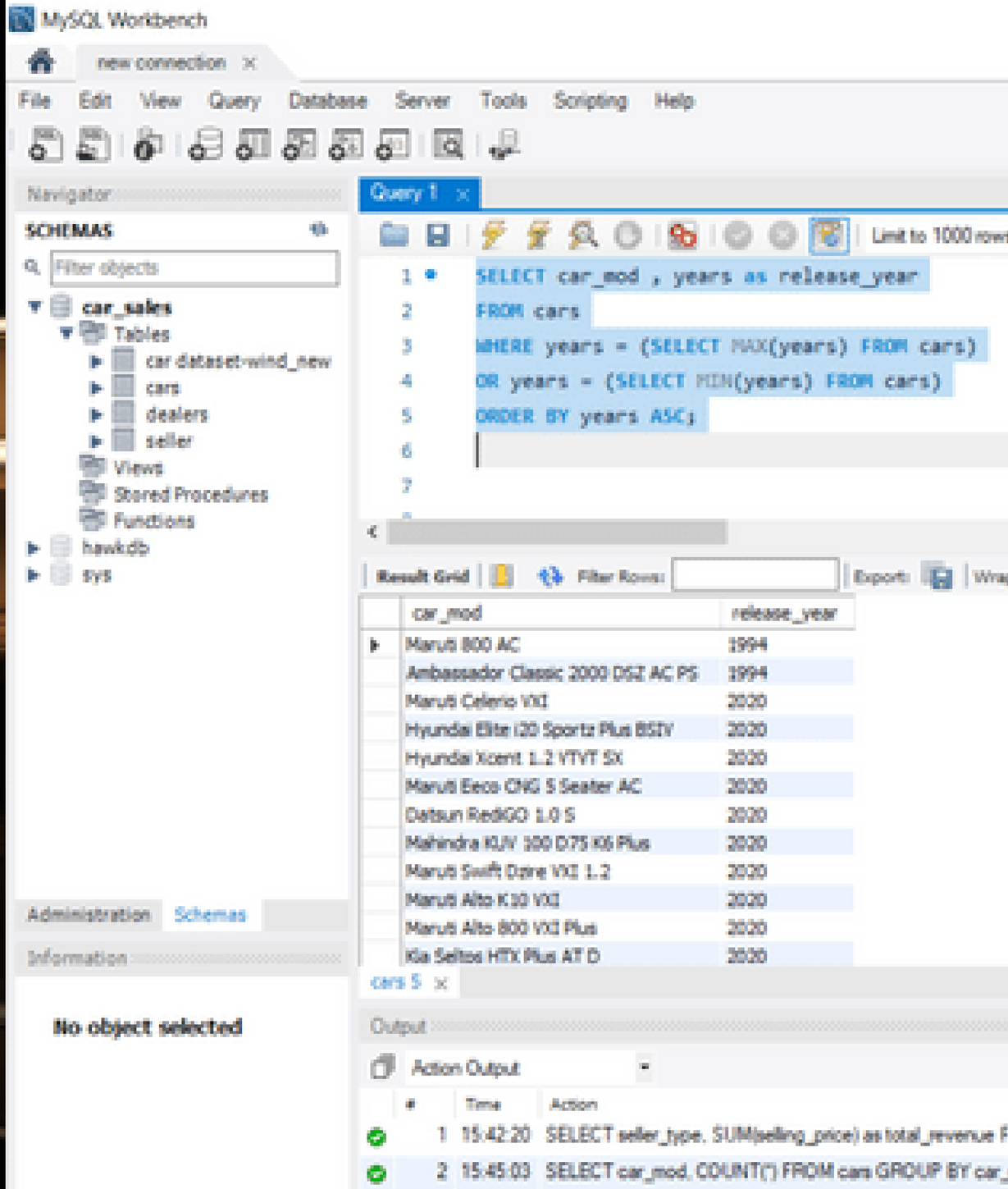
What is the newest car and the oldest car?

We have used the command SELECT car_mod , years as relase_year from table cars , this will show the release year of cars .

We have given the command (SELECT MAX(years) FROM cars),and (SELECT MIN(years) FROM cars) ,this will give the minimum and maximum years attributes .
ORDER BY years ASC; will arrange the outcome data in ascending order .

CONCLUSION - NEWEST CAR - MARUTI CELERIO VXI ,HYUNDAI ELITE I20 SPORTS PLUS BS IV(2020) (+38 more cars in same year, due to no data on exact release date) ETC .

OLDEST CAR IS MARUTI 800 AC (1994) AND AMBASSADOR CLASSIC 2000 DSZ AC PS.



Which is the best car related to price, and low mileage?

1. We have used the command SELECT car_mod, selling_price, mileage
this will select the given attributes from cars table , then we will find the minimum selling price and mileage.

2. ORDER BY selling_price asc, which means we have arranged the selling price in ascending order and mileage

CONCLUSION - MARUTI 800 AC HAS A SELLING OF 29999.00 and mileage 16.1 kmpl as shown in the screenshot. THE OTHER CARS HAS 0.0KMPL MILEAGE WHICH IS UNFEASIBLE .

What is the total revenue from Individual sale?

1. We have given the command SELECT seller_type, SUM(selling_price) as total_revenue ,this will calculate the sum of selling price .

2. With the command LEFT JOIN will join the table seller and cars on Sales_id the WHERE command will give the results for individual seller.

CONCLUSION - INDIVIDUAL SALE = 3,332,153,320.00

QUERY USED

```
ALTER TABLE cars
ADD PRIMARY KEY (sales_id)

ALTER TABLE dealers
ADD FOREIGN KEY (sales_id)
REFERENCES cars(sales_id);

ALTER TABLE seller
ADD FOREIGN KEY (sales_id)
REFERENCES cars(sales_id);
```

01

```
SELECT car_mod, COUNT(*)
FROM cars
GROUP BY car_mod
ORDER BY COUNT(*) DESC
```

02

```
SELECT Dealer , COUNT(dealers.sales_id) as count,
SUM(selling_price) as total_sales
FROM dealers LEFT JOIN cars on dealers.sales_id =
cars.sales_id
GROUP BY Dealer
ORDER BY total_sales desc;
```

03

```
SELECT car_mod, AVG(selling_price) as average_price
FROM cars
GROUP BY car_mod
ORDER BY average_price DESC;
```

04

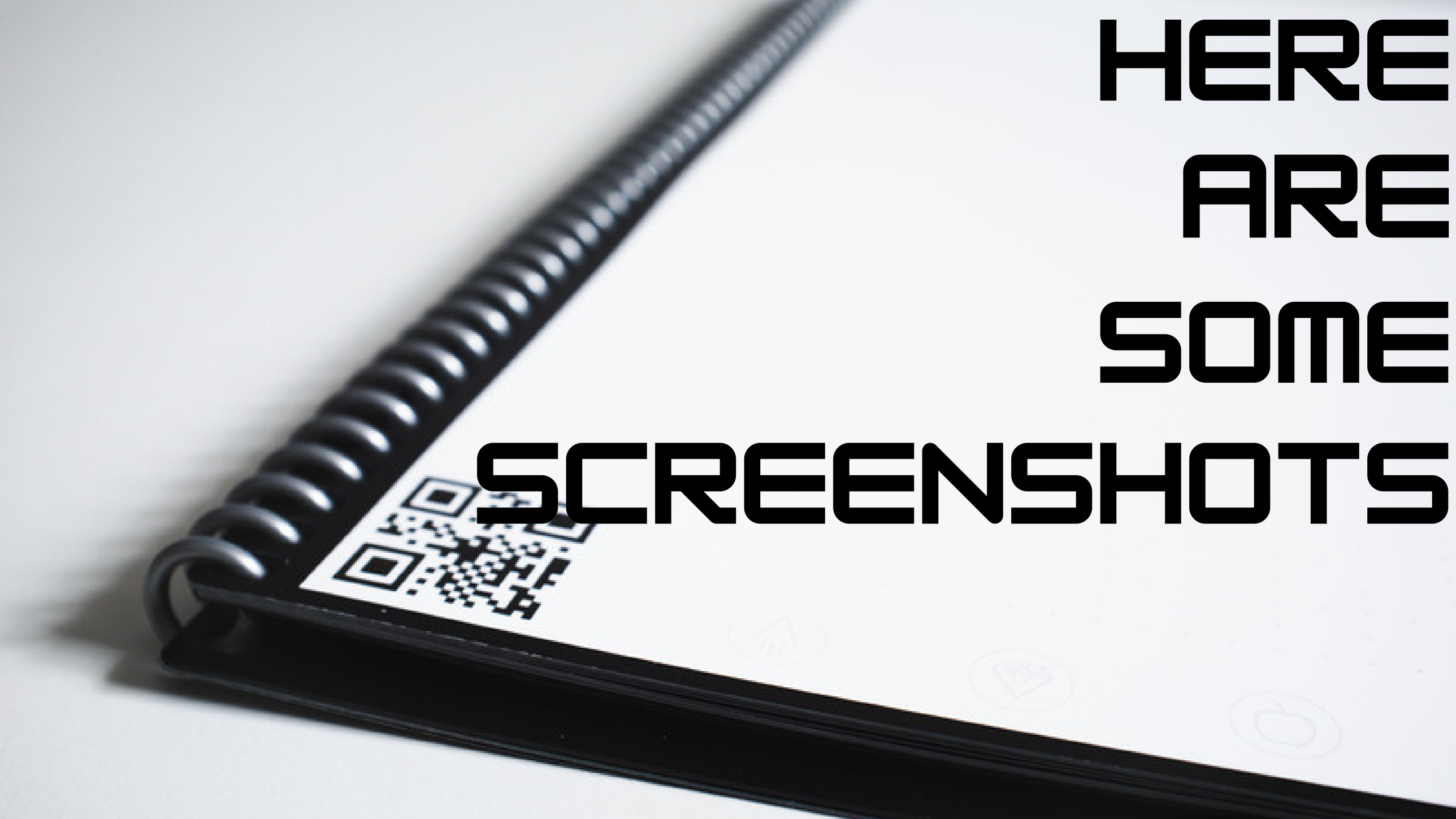
```
SELECT car_mod , years as release_year
FROM cars
WHERE years = (SELECT MAX(years) FROM cars)
OR years = (SELECT MIN(years) FROM cars)
ORDER BY years ASC;
```

05

```
SELECT car_mod, selling_price, mileage
FROM cars WHERE selling_price = (SELECT MIN(selling_price)
FROM cars) OR mileage = (SELECT MIN(mileage) FROM cars)
ORDER BY selling_price asc, mileage asc;
```

06

```
SELECT seller_type, SUM(selling_price) as total_reven
FROM seller LEFT JOIN cars ON
seller.sales_id = cars.sales_id
WHERE seller_type = 'Individual';
```


HERE
ARE
SOME
SCREENSHOTS

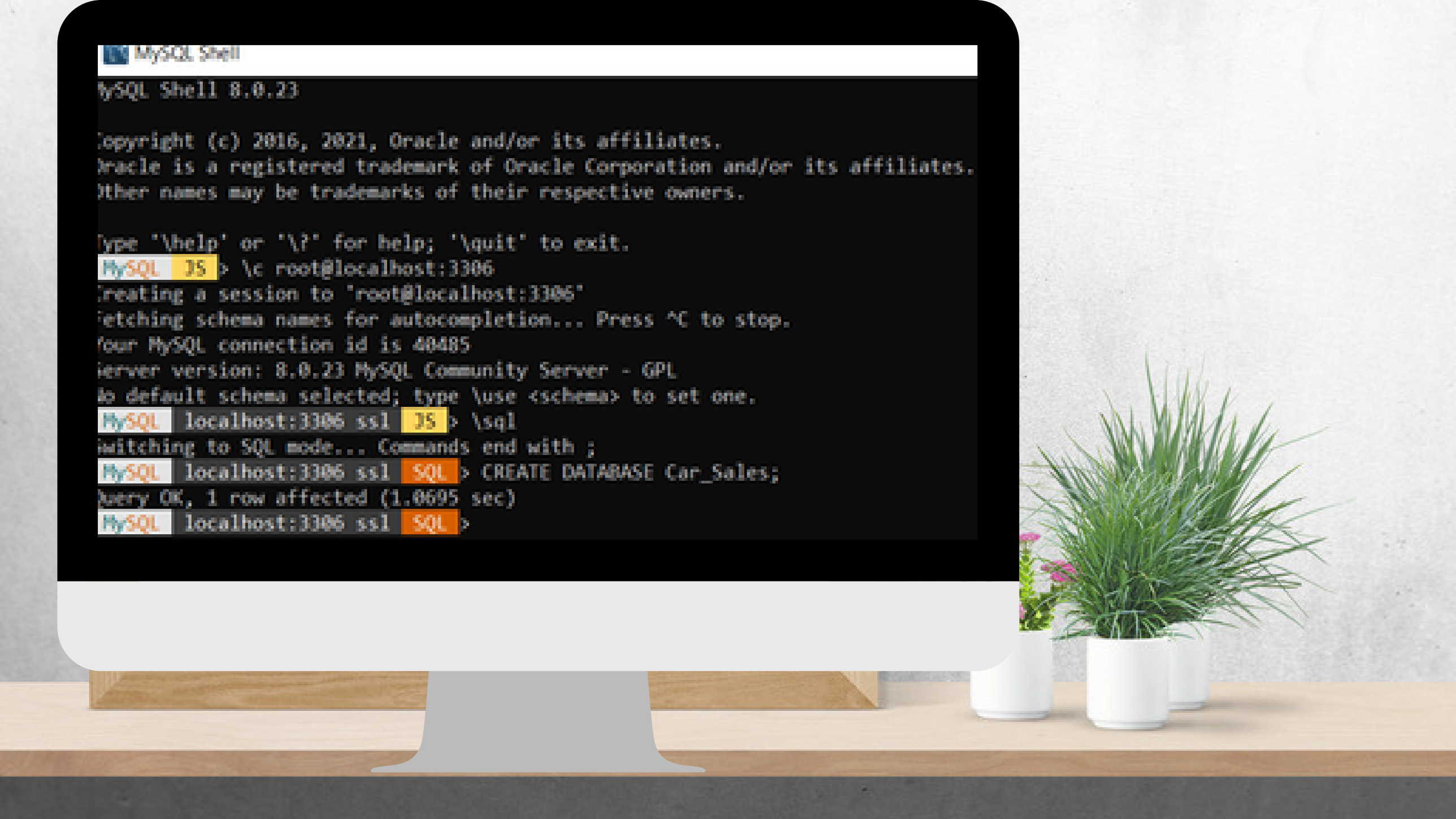

MySQL Shell
MySQL Shell 8.0.23
Copyright (c) 2016, 2021, Oracle and/or its affiliates.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates.
Other names may be trademarks of their respective owners.
Type '\help' or '\?' for help; '\quit' to exit.
MySQL JS > \c root@localhost:3306
Creating a session to 'root@localhost:3306'
Fetching schema names for autocompletion... Press ^C to stop.
Your MySQL connection id is 40485
Server version: 8.0.23 MySQL Community Server - GPL
No default schema selected; type \use <schema> to set one.
MySQL localhost:3306 ssl JS > \sql
Switching to SQL mode... Commands end with ;
MySQL localhost:3306 ssl SQL > CREATE DATABASE Car_Sales;
Query OK, 1 row affected (1.0695 sec)
MySQL localhost:3306 ssl SQL >

TABLE COMMANDS SCREENSHOT

MySQL Workbench

Local instance MySQL80

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- firstdb
- projectdb
  - Tables
    - cars
    - dealers
    - main
    - seller
  - Views
  - Stored Procedures
  - Functions
- sys

Administration Schemas

Information

Table: dealers

Columns:
dealer char(30)
sales_id int

Object Info Session

Query 1 | SQL File 3* | SQL File 4* | main | cars | seller | dealers

Limit to 10000 rows

```
18  create table Cars(
19  sales_id int,
20  name varchar(100),
21  year int(10),
22  km_driven int(10),
23  selling_price char(10),
24  fuel varchar(20),
25  transmission char(30),
26  mileage char(20),
27  engine char(30),
28  max_power char(30),
29  torque char(100),
30  seats int (3)
31  );
32
33  create table Seller(
34  owner varchar(20),
35  seller_type varchar(20),
36  sales_id int(5)
37  );
38
39  create table dealers(
40  dealer char(30),
```

SQL Additions

Jump to

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

TO BE CONTINUE

Context Help Snippets

Output

Action Output

#	Time	Action	Message	Duration / Fetch
34	22:58:22	PREPARE stmt FROM 'INSERT INTO `projectdb`.`dealers` (`dealer`,`sales_id`) VALUES(?,?)'	OK	0.000 sec
35	22:58:28	DEALLOCATE PREPARE stmt	OK	0.000 sec
36	22:58:36	SELECT * FROM projectdb.dealers LIMIT 0, 10000	1107 row(s) returned	0.015 sec / 0.000 sec

Search

ENG 23:38 22-05-2021

MySQL Workbench

Local instance MySQL80

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- firstdb
- projectdb
  - Tables
    - cars
    - dealers
    - main
    - seller
  - Views
  - Stored Procedures
  - Functions
- sys

Administration Schemas

Information

Table: dealers

Columns:
dealer char(30)
sales_id int

Object Info Session

Query 1 SQL File 3* SQL File 4* main cars seller dealers

Limit to 10000 rows

```
22    km_driven int(10),
23    selling_price char(10),
24    fuel varchar(20),
25    transmission char(30),
26    mileage char(20),
27    engine char(30),
28    max_power char(30),
29    torque char(100),
30    seats int (3)
31    );
32
33    create table Seller(
34    owner varchar(20),
35    seller_type varchar(20),
36    sales_id int(5)
37    );
38
39    create table dealers(
40    dealer char(30),
41    sales_id int
42    );
43
44
```

SQL Additions

Jump to

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Context Help Snippets

Output

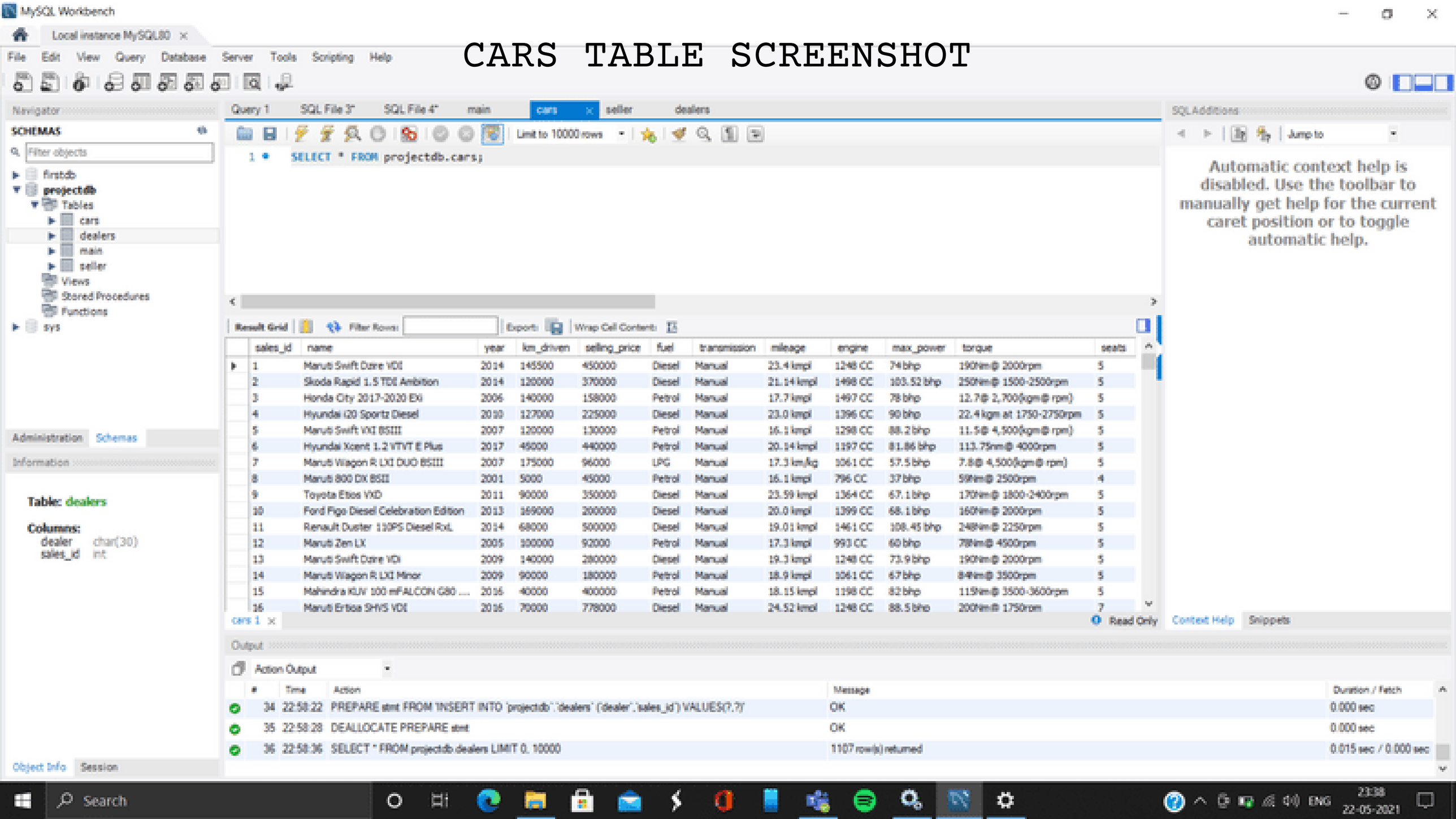
Action Output

#	Time	Action	Message	Duration / Fetch
34	22:58:22	PREPARE stmt FROM 'INSERT INTO `projectdb`.`dealers` (`dealer`,`sales_id`) VALUES(?,?)'	OK	0.000 sec
35	22:58:28	DEALLOCATE PREPARE stmt	OK	0.000 sec
36	22:58:36	SELECT * FROM projectdb.dealers LIMIT 0, 10000	1107 row(s) returned	0.015 sec / 0.000 sec

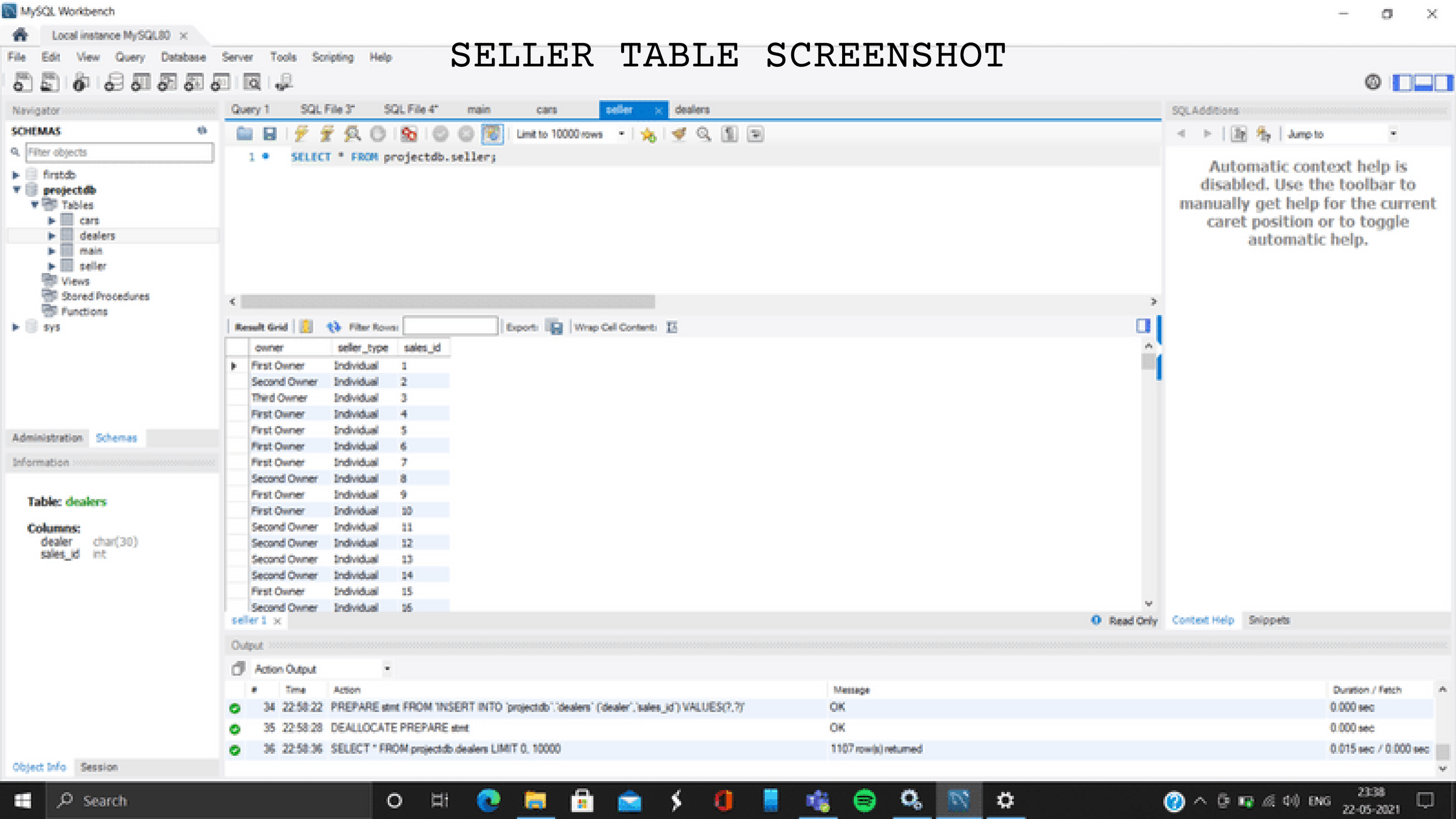
Search

ENG 23:38 22-05-2021

CARS TABLE SCREENSHOT



SELLER TABLE SCREENSHOT



DEALER TABLE SCREENSHOT