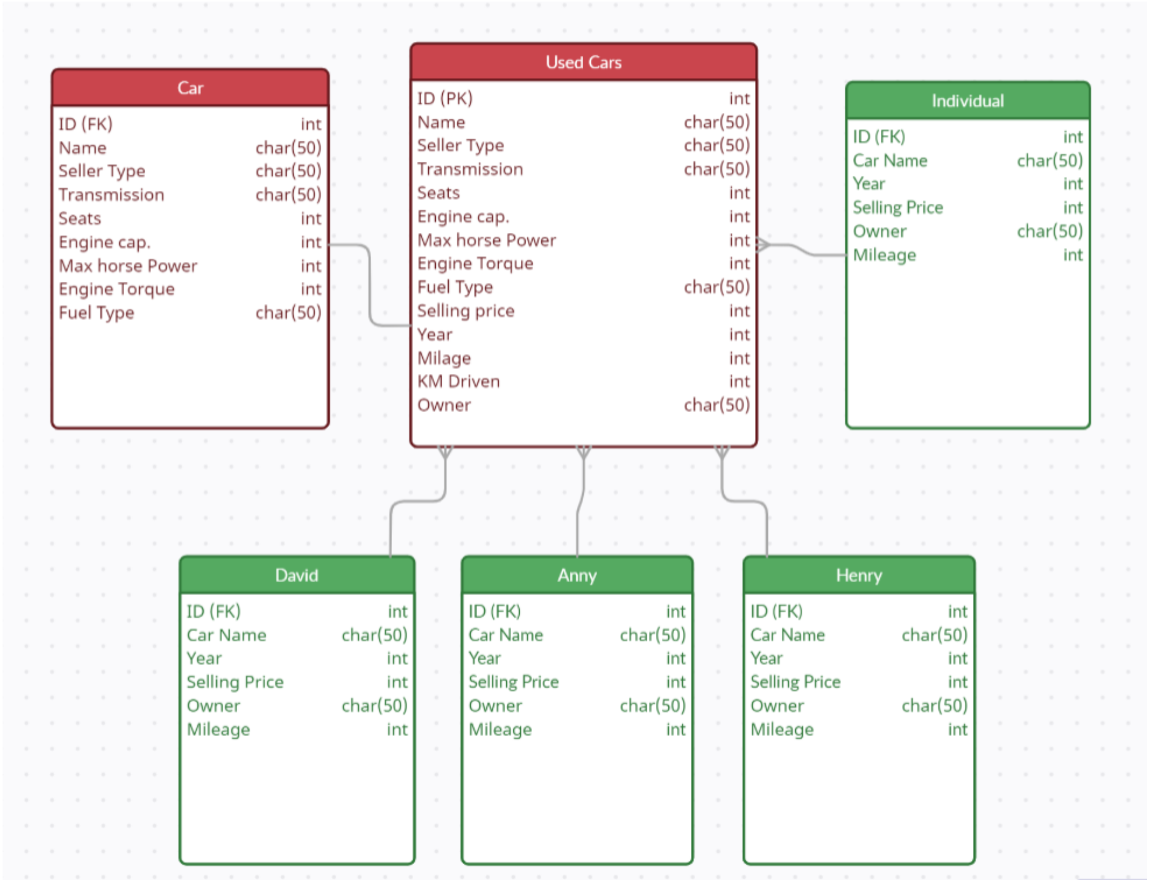
University of Europe for Applied Science

Big Data and Analytics: Final Project
Part One: Creating Relational Database
Case Study: Used Car Agency
REPORT

Supervisor: Dr. Rand Kouatly
Names of the students: Kunwardeep Singh, Michal Gorecki,
Mohab Mousa, Nurillokh Abdurasulov, Laura-Sophia Ließ
Study program: Digital Business and Data Science

1. Explanation of the database design with the ER diagram & the SQL code

We started by creating a database called “CAR_AGENCY”.
Our main table is the “Used Cars”-table, filled with our primary key “ID” as integer and all the information given, which are name, seller type, transmission, seats, engine capacity, maximum horsepower, engine torque, fuel type, selling price, year, mileage, kilometers driven and owner. The name, seller type, transmission, fuel type and owner as maximum 50 characters and the other information are all given as integers. The table has a one-to-one connection to the “Car”-table.
Our table “Car” has the technical information about the car. We added the primary key “ID”, and the following information: name, seller type, transmission, seats, engine capacity, maximum horsepower, engine torque and fuel type.
From the “Used Cars”-table we made 4 other connections to the tables “Individual”, “David”, “Anny” and “Henry”. They are all many-to-one connections, because many cars can be sold by just one dealer. This has the advantage of a simple overview about who sold which and how many cars. All of those 4 tables have the same structure. It starts with the foreign key “ID”, which is the primary key in the “Used Cars”-table. Moreover the tables contain the information which are determining the value of the car and information about the selling which are car name, year, selling price, owner and mileage.



Car	
ID (FK)	int
Name	char(50)
Seller Type	char(50)
Transmission	char(50)
Seats	int
Engine cap.	int
Max horse Power	int
Engine Torque	int
Fuel Type	char(50)

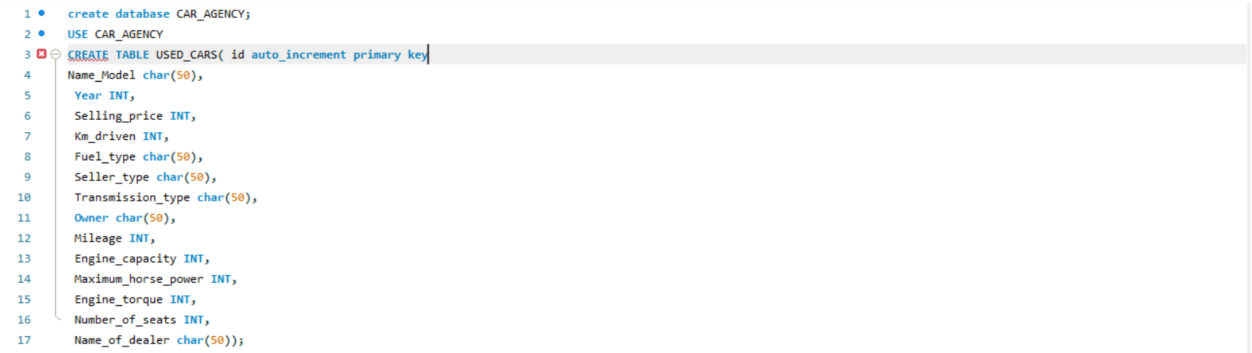
Used Cars	
ID (PK)	int
Name	char(50)
Seller Type	char(50)
Transmission	char(50)
Seats	int
Engine cap.	int
Max horse Power	int
Engine Torque	int
Fuel Type	char(50)
Selling price	int
Year	int
Milage	int
KM Driven	int
Owner	char(50)

Individual	
ID (FK)	int
Car Name	char(50)
Year	int
Selling Price	int
Owner	char(50)
Mileage	int

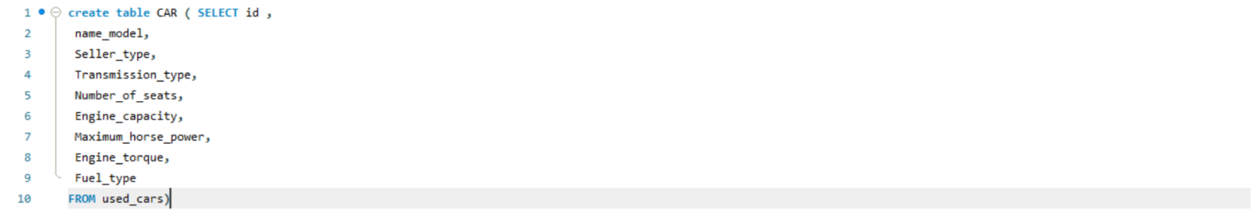
David	
ID (FK)	int
Car Name	char(50)
Year	int
Selling Price	int
Owner	char(50)
Mileage	int

Anny	
ID (FK)	int
Car Name	char(50)
Year	int
Selling Price	int
Owner	char(50)
Mileage	int

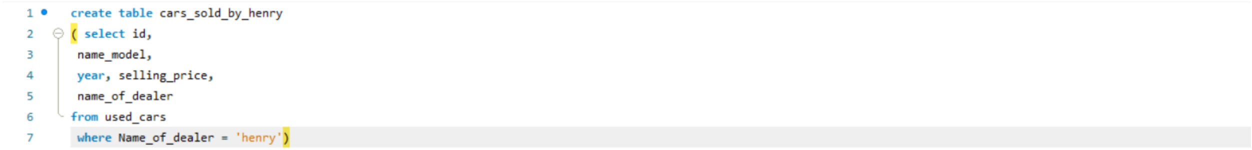
Henry	
ID (FK)	int
Car Name	char(50)
Year	int
Selling Price	int
Owner	char(50)
Mileage	int



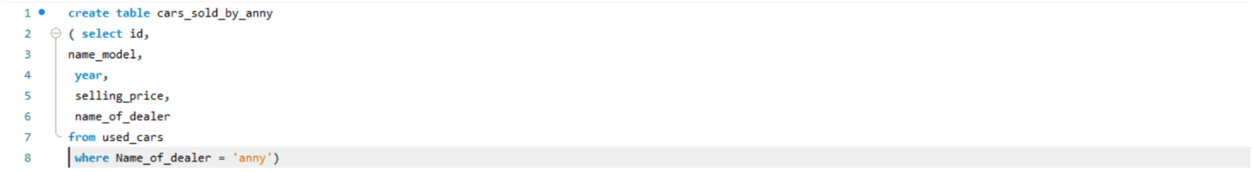
```sql
create database CAR_AGENCY;
USE CAR_AGENCY
CREATE TABLE USED_CARS( id auto_increment primary key
Name_Model char(50),
 Year INT,
 Selling_price INT,
 Km_driven INT,
 Fuel_type char(50),
 Seller_type char(50),
 Transmission_type char(50),
 Owner char(50),
 Mileage INT,
 Engine_capacity INT,
 Maximum_horse_power INT,
 Engine_torque INT,
 Number_of_seats INT,
 Name_of_dealer char(50));
```



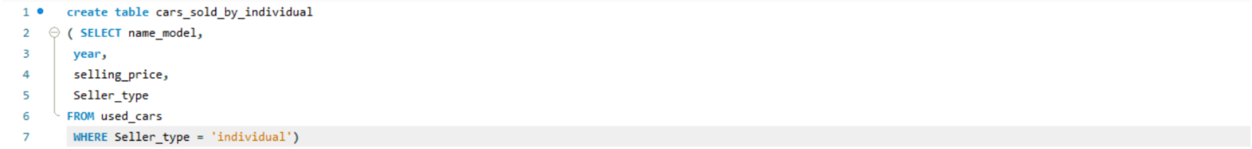
```sql
create table CAR ( SELECT id ,
 name_model,
 Seller_type,
 Transmission_type,
 Number_of_seats,
 Engine_capacity,
 Maximum_horse_power,
 Engine_torque,
 Fuel_type
FROM used_cars)
```



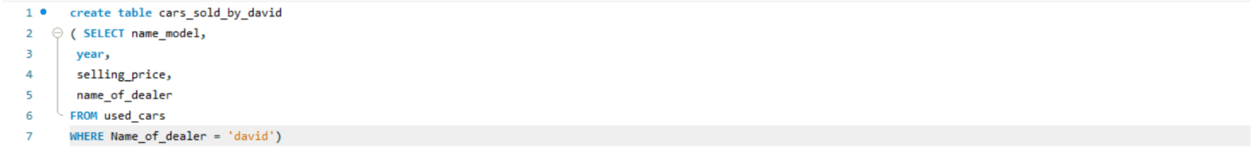
```sql
create table cars_sold_by_henry
( select id,
 name_model,
 year, selling_price,
 name_of_dealer
from used_cars
 where Name_of_dealer = 'henry')
```



```sql
create table cars_sold_by_anny
( select id,
name_model,
 year,
 selling_price,
 name_of_dealer
from used_cars
 where Name_of_dealer = 'anny')
```

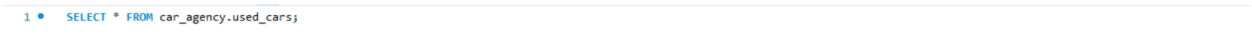


```sql
create table cars_sold_by_individual
( SELECT name_model,
 year,
 selling_price,
 Seller_type
FROM used_cars
 WHERE Seller_type = 'individual')
```

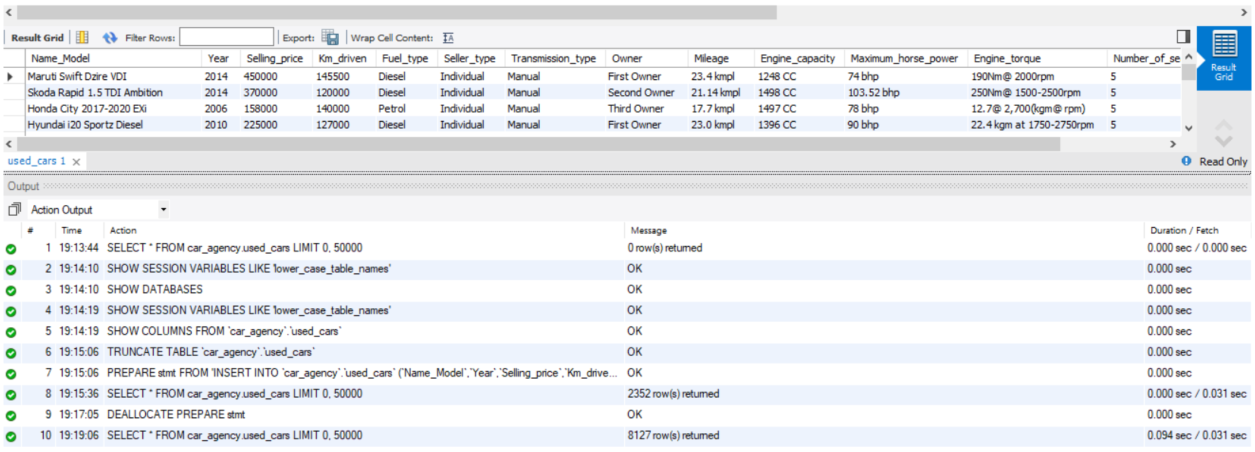


```sql
create table cars_sold_by_david
( SELECT name_model,
 year,
 selling_price,
 name_of_dealer
FROM used_cars
WHERE Name_of_dealer = 'david')
```

2. Evidence that database is created (screenshot)



```sql
SELECT * FROM car_agency.used_cars;
```



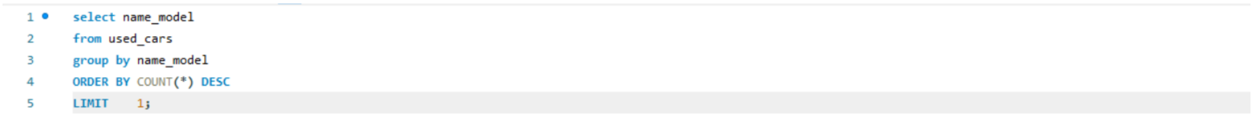
Name_Model	Year	Selling_price	Km_driven	Fuel_type	Seller_type	Transmission_type	Owner	Mileage	Engine_capacity	Maximum_horse_power	Engine_torque	Number_of_se
Maruti Swift Dzire VDI	2014	450000	145500	Diesel	Individual	Manual	First Owner	23.4 kmpl	1248 CC	74 bhp	190Nm@ 2000rpm	5
Skoda Rapid 1.5 TDI Ambition	2014	370000	120000	Diesel	Individual	Manual	Second Owner	21.14 kmpl	1498 CC	103.52 bhp	250Nm@ 1500-2500rpm	5
Honda City 2017-2020 EXi	2006	158000	140000	Petrol	Individual	Manual	Third Owner	17.7 kmpl	1497 CC	78 bhp	12.7@ 2,700(kgm@ rpm)	5
Hyundai i20 Sportz Diesel	2010	225000	127000	Diesel	Individual	Manual	First Owner	23.0 kmpl	1396 CC	90 bhp	22.4 kgm at 1750-2750rpm	5

#	Time	Action	Message	Duration / Fetch
1	19:13:44	SELECT * FROM car_agency.used_cars LIMIT 0, 50000	0 row(s) returned	0.000 sec / 0.000 sec
2	19:14:10	SHOW SESSION VARIABLES LIKE 'lower_case_table_names'	OK	0.000 sec
3	19:14:10	SHOW DATABASES	OK	0.000 sec
4	19:14:19	SHOW SESSION VARIABLES LIKE 'lower_case_table_names'	OK	0.000 sec
5	19:14:19	SHOW COLUMNS FROM `car_agency`.`used_cars`	OK	0.000 sec
6	19:15:06	TRUNCATE TABLE `car_agency`.`used_cars`	OK	0.000 sec
7	19:15:06	PREPARE stmt FROM 'INSERT INTO `car_agency`.`used_cars` (`Name_Model`,`Year`,`Selling_price`,`Km_drive...	OK	0.000 sec
8	19:15:36	SELECT * FROM car_agency.used_cars LIMIT 0, 50000	2352 row(s) returned	0.000 sec / 0.031 sec
9	19:17:05	DEALLOCATE PREPARE stmt	OK	0.000 sec
10	19:19:06	SELECT * FROM car_agency.used_cars LIMIT 0, 50000	8127 row(s) returned	0.094 sec / 0.031 sec

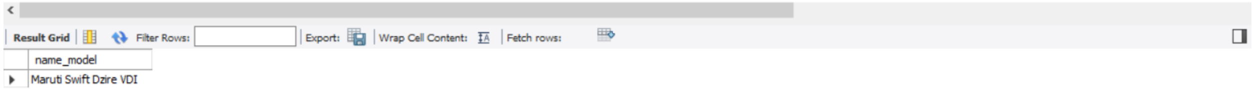
3. Q&A

- **Find out which car models are sold the most?**

Answer: Maruti Swift Dzire VDI



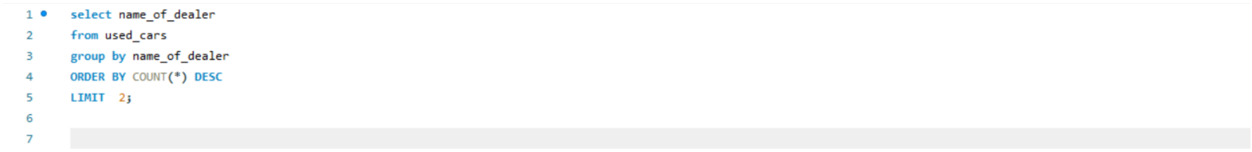
```sql
select name_model
from used_cars
group by name_model
ORDER BY COUNT(*) DESC
LIMIT    1;
```



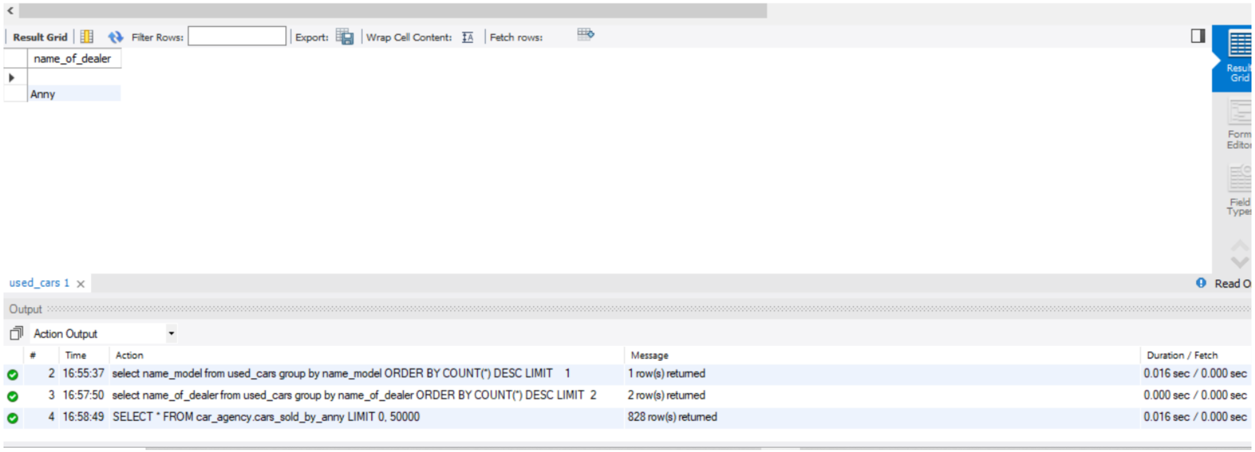
name_model
Maruti Swift Dzire VDI

- **Which dealer sold the most cars, and what is his/her total sales?**

Answer: Anny with 828 cars (the rows at the bottom)



```sql
select name_of_dealer
from used_cars
group by name_of_dealer
ORDER BY COUNT(*) DESC
LIMIT  2;
```



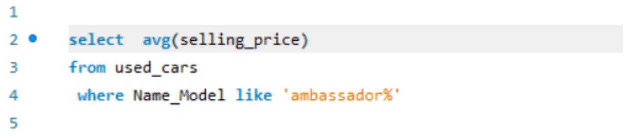
name_of_dealer
Anny

```
#   Time      Action                                                                              Message               Duration / Fetch
2   16:55:37  select name_model from used_cars group by name_model ORDER BY COUNT(*) DESC LIMIT   1   1 row(s) returned     0.016 sec / 0.000 sec
3   16:57:50  select name_of_dealer from used_cars group by name_of_dealer ORDER BY COUNT(*) DESC LIMIT  2   2 row(s) returned     0.000 sec / 0.000 sec
4   16:58:49  SELECT * FROM car_agency.cars_sold_by_anny LIMIT 0, 50000                          828 row(s) returned   0.016 sec / 0.000 sec
```

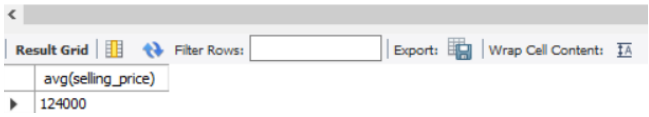
- **What is the average price for each car model?**

Answer:

Ambassador:

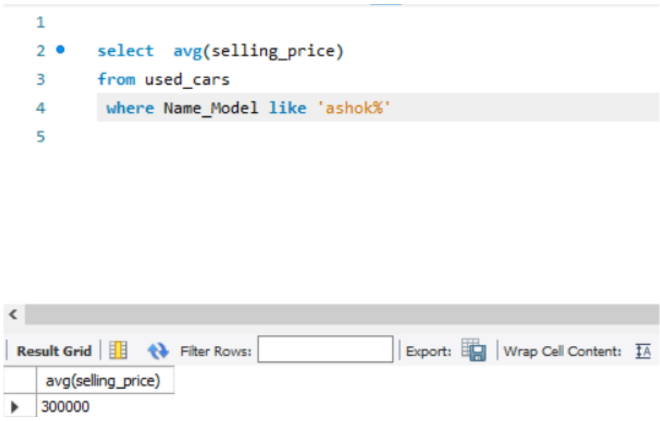


```sql

select  avg(selling_price)
from used_cars
 where Name_Model like 'ambassador%'

```



avg(selling_price)
124000

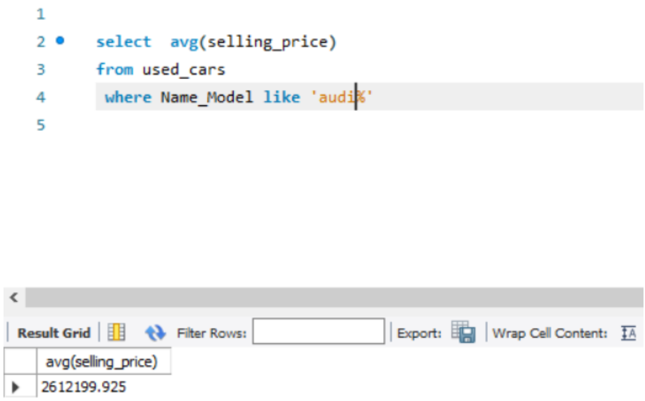
Ashok:



```
select  avg(selling_price)
from used_cars
 where Name_Model like 'ashok%'
```

```
avg(selling_price)
300000
```

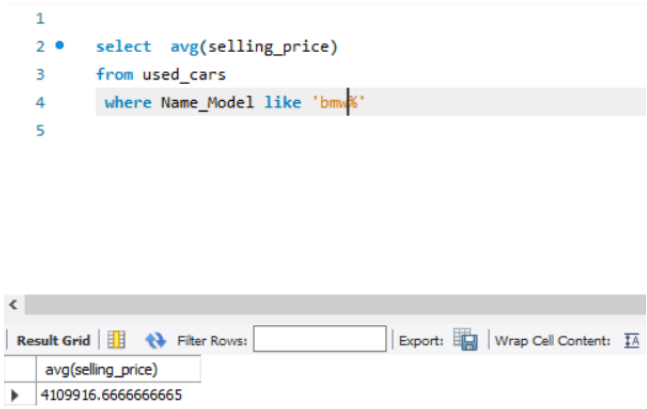
Audi:



```
select  avg(selling_price)
from used_cars
 where Name_Model like 'audi%'
```

```
avg(selling_price)
2612199.925
```

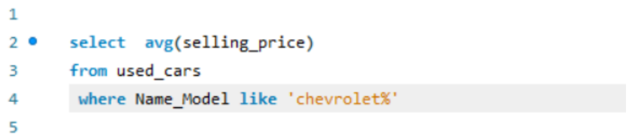
BMW:



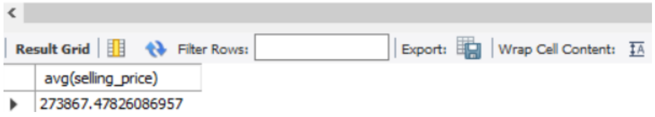
```
select  avg(selling_price)
from used_cars
 where Name_Model like 'bmw%'
```

```
avg(selling_price)
4109916.6666666665
```

Chevrolet:

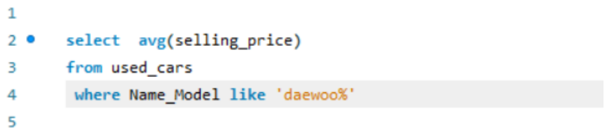


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'chevrolet%'
```

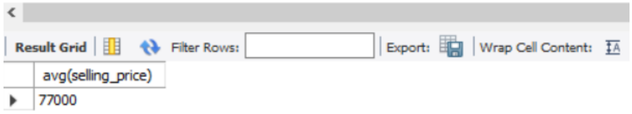


```
avg(selling_price)
273867.47826086957
```

Daewoo:

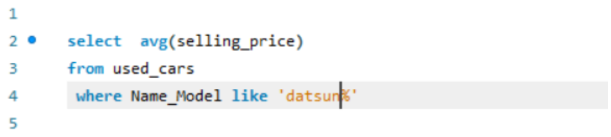


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'daewoo%'
```

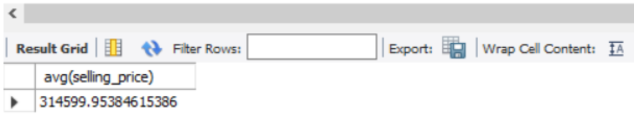


```
avg(selling_price)
77000
```

Datsun:

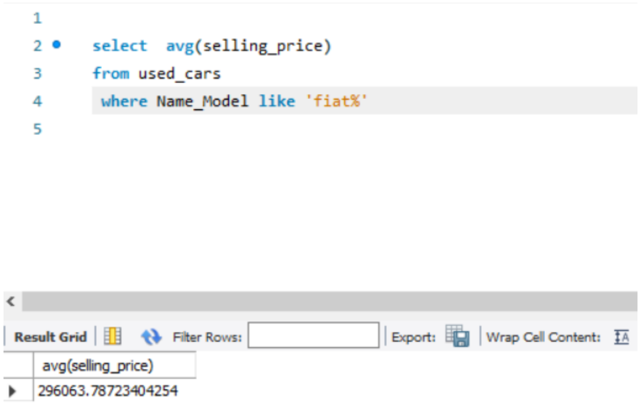


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'datsun%'
```



```
avg(selling_price)
314599.95384615386
```

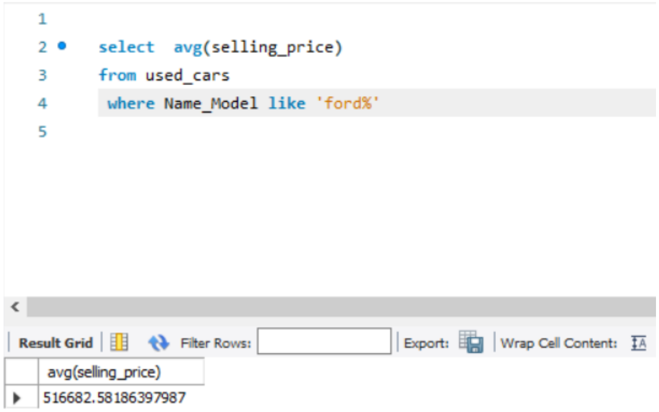
Fiat:



```
select  avg(selling_price)
from used_cars
 where Name_Model like 'fiat%'
```

```
avg(selling_price)
296063.78723404254
```

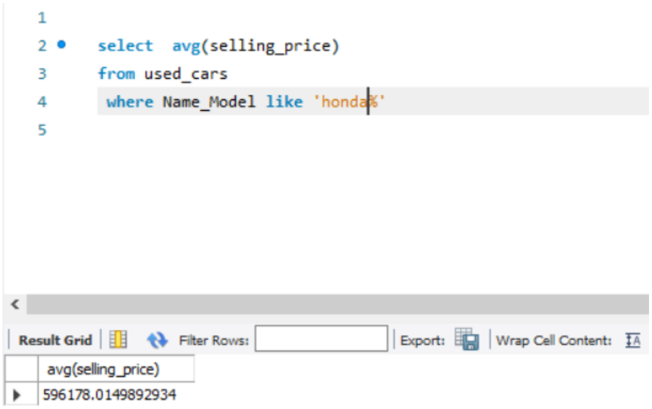
Ford:



```
select  avg(selling_price)
from used_cars
 where Name_Model like 'ford%'
```

```
avg(selling_price)
516682.58186397987
```

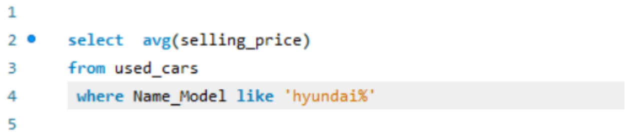
Honda:



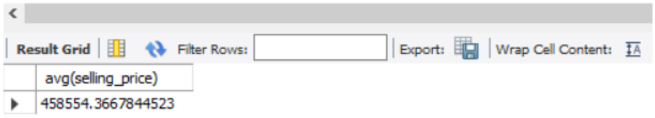
```
select  avg(selling_price)
from used_cars
 where Name_Model like 'honda%'
```

```
avg(selling_price)
596178.0149892934
```

Hyundai:

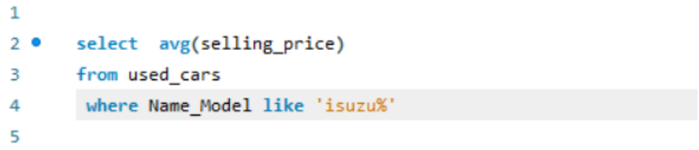


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'hyundai%'
```

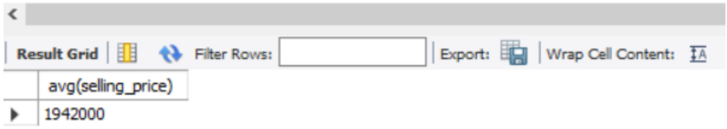


```
avg(selling_price)
458554.3667844523
```

Isuzu:

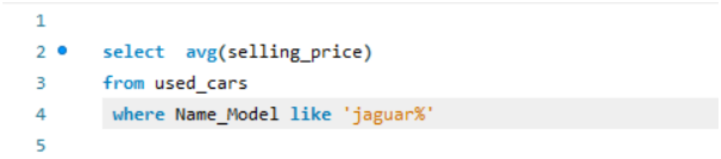


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'isuzu%'
```

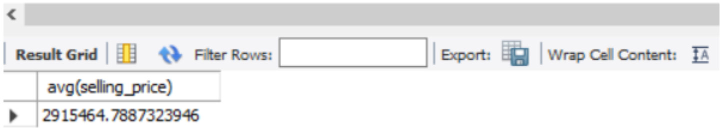


```
avg(selling_price)
1942000
```

Jaguar:

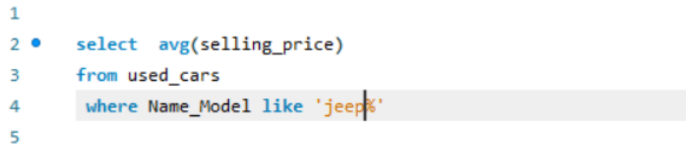


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'jaguar%'
```

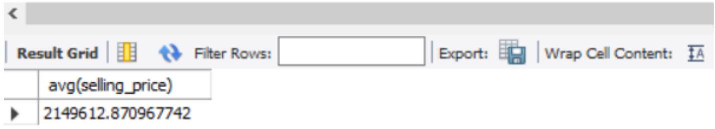


```
avg(selling_price)
2915464.7887323946
```

Jeep:

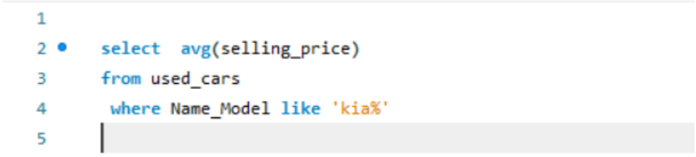


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'jeep%'
```

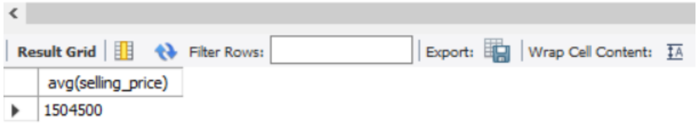


```
avg(selling_price)
2149612.870967742
```

Kia:

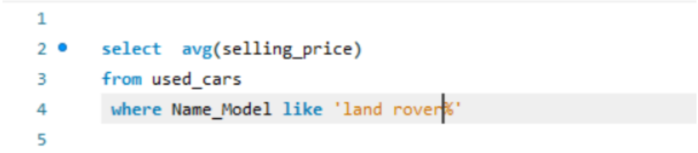


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'kia%'
```

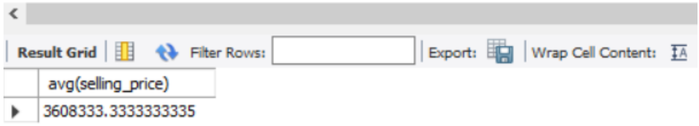


```
avg(selling_price)
1504500
```

Land Rover:

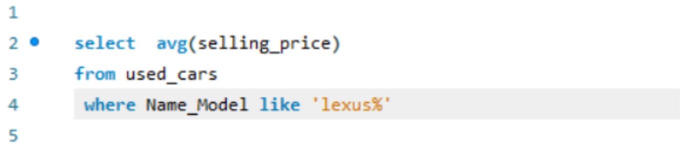


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'land rover%'
```

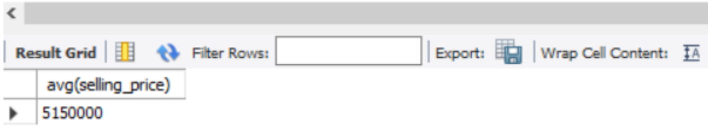


```
avg(selling_price)
3608333.3333333335
```

Lexus:

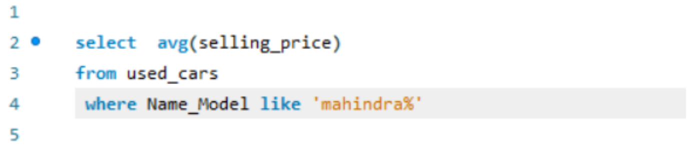


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'lexus%'
```

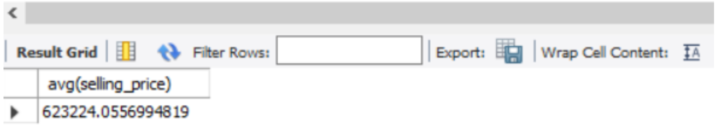


```
avg(selling_price)
5150000
```

Mahindra:

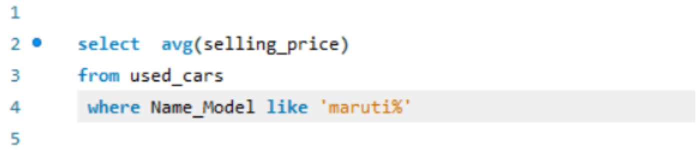


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'mahindra%'
```

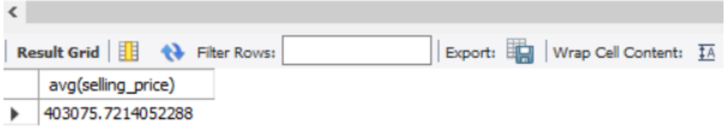


```
avg(selling_price)
623224.0556994819
```

Maruti:

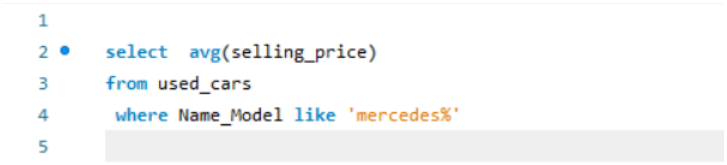


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'maruti%'
```

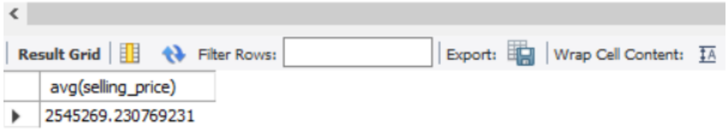


```
avg(selling_price)
403075.7214052288
```

Mercedes:

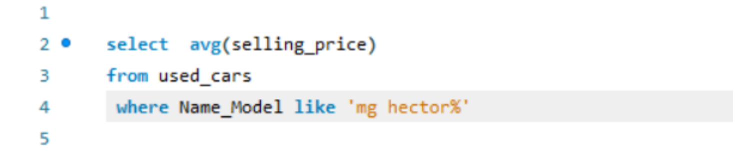


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'mercedes%'
```

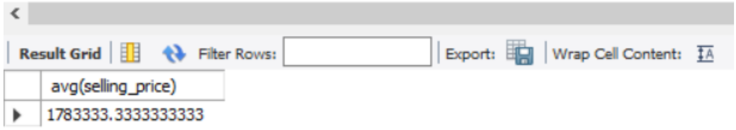


```
avg(selling_price)
2545269.230769231
```

MG Hector:

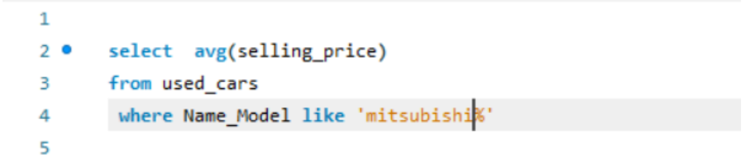


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'mg hector%'
```

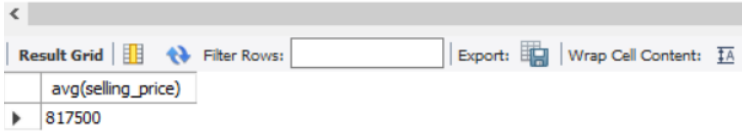


```
avg(selling_price)
1783333.3333333333
```

Mitsubishi:

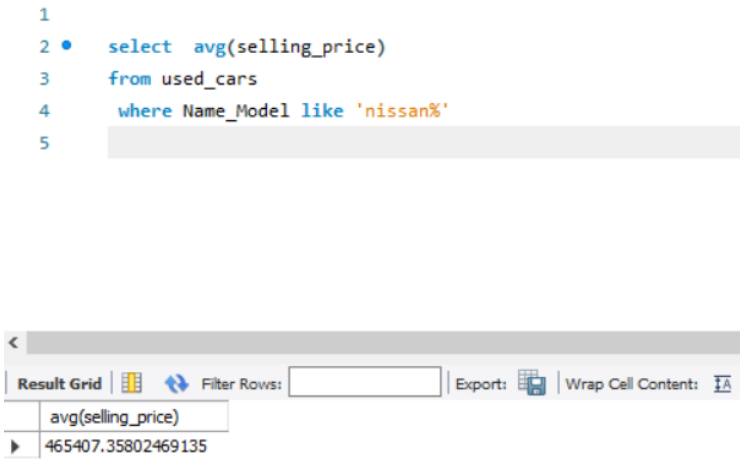


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'mitsubishi%'
```



```
avg(selling_price)
817500
```

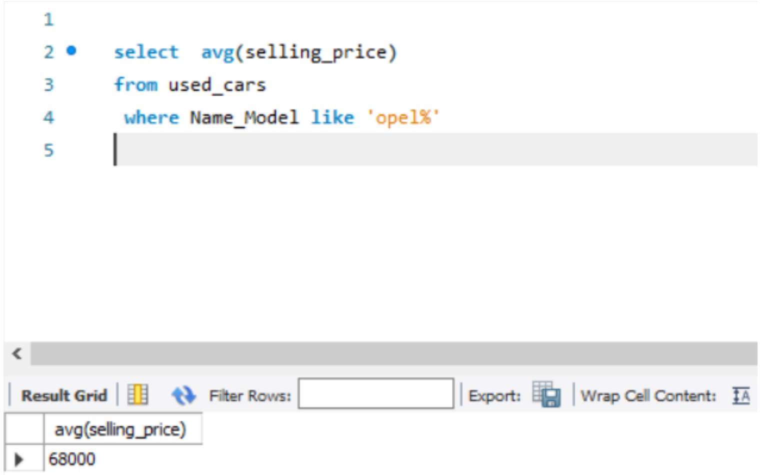
Nissan:



```

select  avg(selling_price)
from used_cars
 where Name_Model like 'nissan%'

```

```
avg(selling_price)
465407.35802469135
```

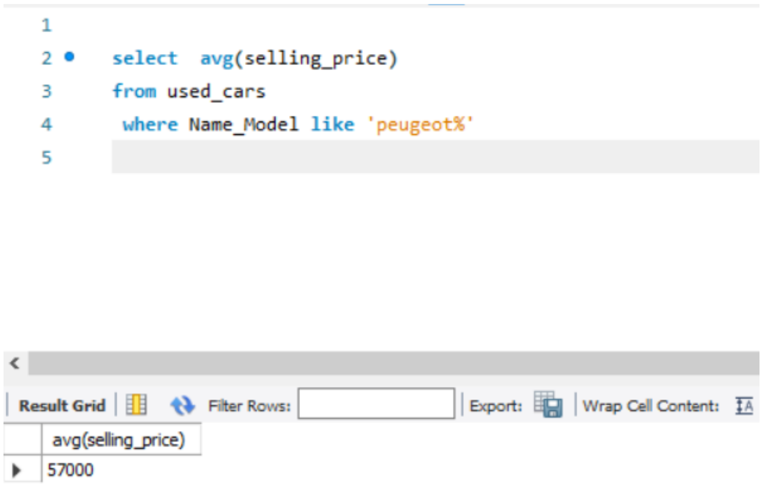
Opel:



```

select  avg(selling_price)
from used_cars
 where Name_Model like 'opel%'

```

```
avg(selling_price)
68000
```

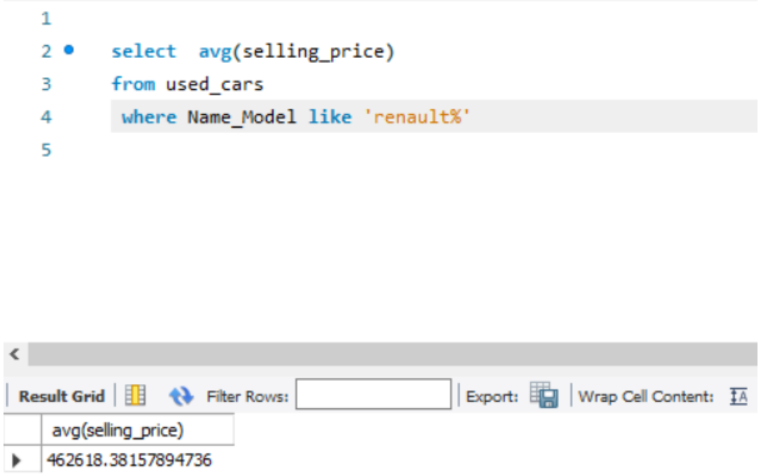
Peugeot:



```

select  avg(selling_price)
from used_cars
 where Name_Model like 'peugeot%'

```

```
avg(selling_price)
57000
```

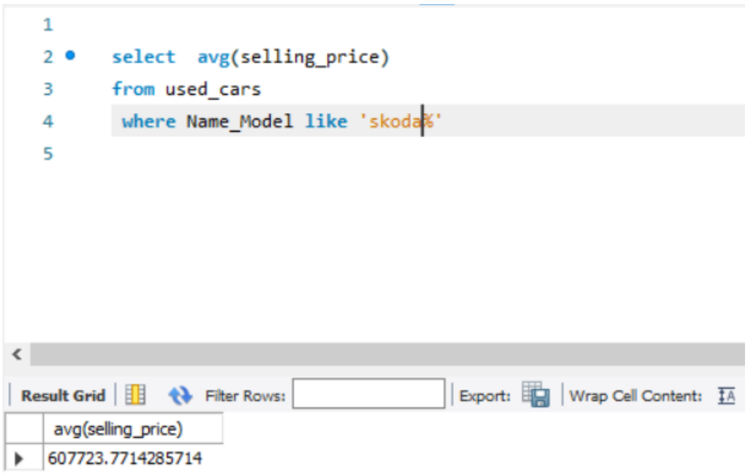
Renault:



```sql

select  avg(selling_price)
from used_cars
 where Name_Model like 'renault%'

```

```
avg(selling_price)
462618.38157894736
```

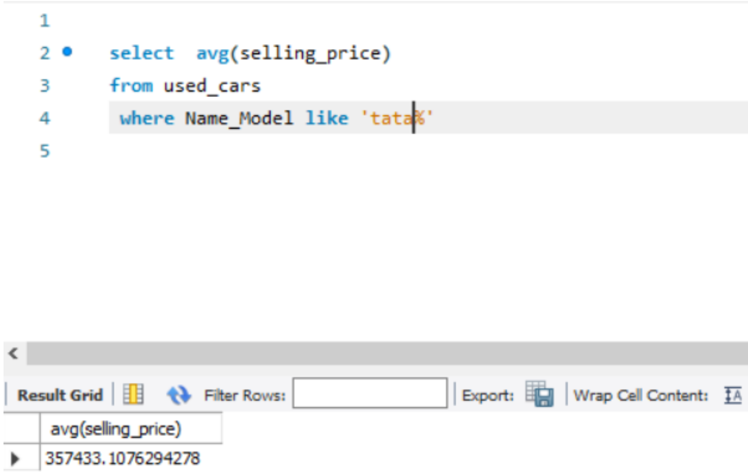
Skoda:



```sql

select  avg(selling_price)
from used_cars
 where Name_Model like 'skoda%'

```

```
avg(selling_price)
607723.7714285714
```

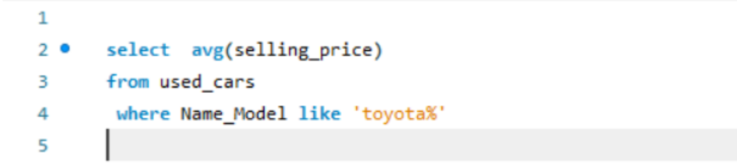
Tata:



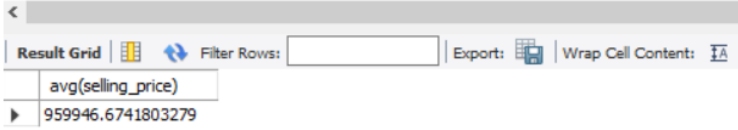
```sql

select  avg(selling_price)
from used_cars
 where Name_Model like 'tata%'

```

```
avg(selling_price)
357433.1076294278
```

Toyota:

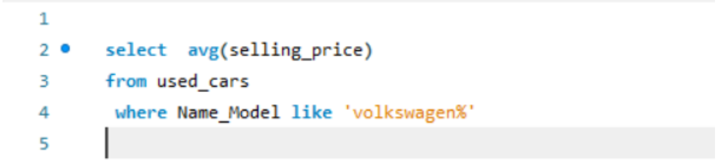


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'toyota%'
```

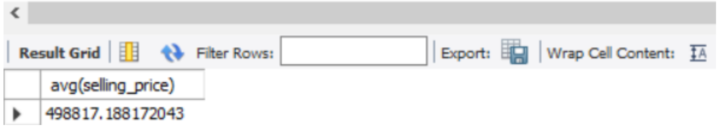


```
avg(selling_price)
959946.6741803279
```

Volkswagen:

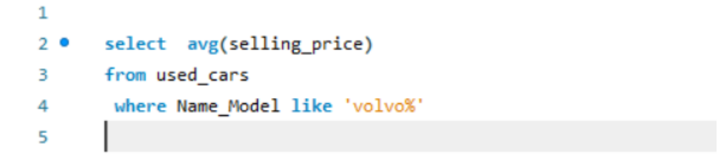


```
select  avg(selling_price)
from used_cars
 where Name_Model like 'volkswagen%'
```

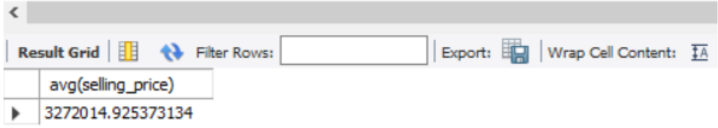


```
avg(selling_price)
498817.188172043
```

Volvo:



```
select  avg(selling_price)
from used_cars
 where Name_Model like 'volvo%'
```

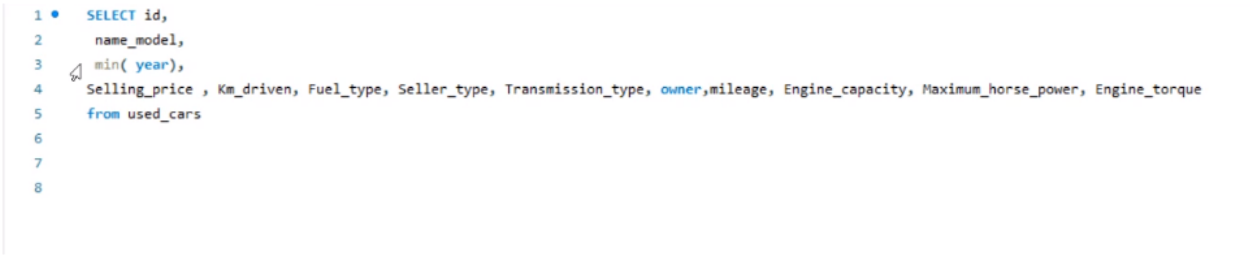


```
avg(selling_price)
3272014.925373134
```

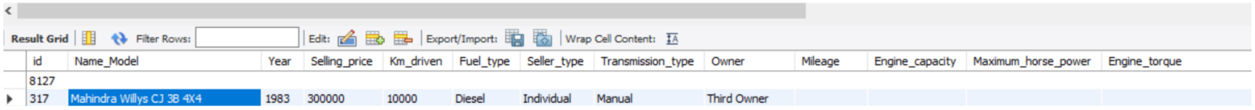
- **What is the newest and the oldest car?**

Answers:

Oldest: Mahindra Willys CJ 38 4X4; year: 1983

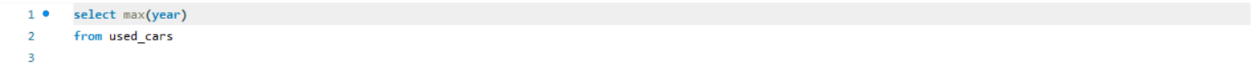


```sql
SELECT id,
 name_model,
 min( year),
Selling_price , Km_driven, Fuel_type, Seller_type, Transmission_type, owner,mileage, Engine_capacity, Maximum_horse_power, Engine_torque
from used_cars
```

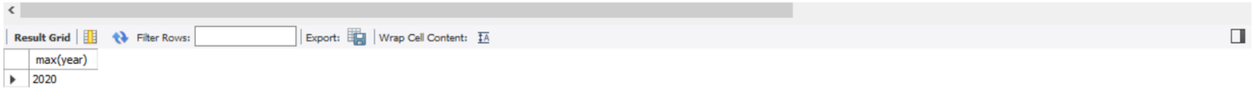


id	Name_Model	Year	Selling_price	Km_driven	Fuel_type	Seller_type	Transmission_type	Owner	Mileage	Engine_capacity	Maximum_horse_power	Engine_torque
8127												
317	Mahindra Willys CJ 3B 4X4	1983	300000	10000	Diesel	Individual	Manual	Third Owner				

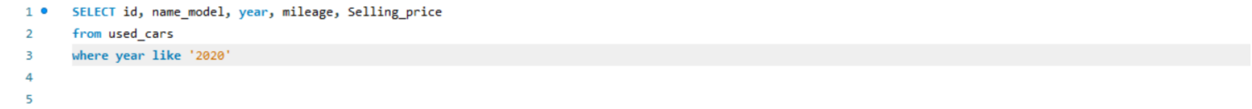
Newest: Multiple cars are from the same year; 2020



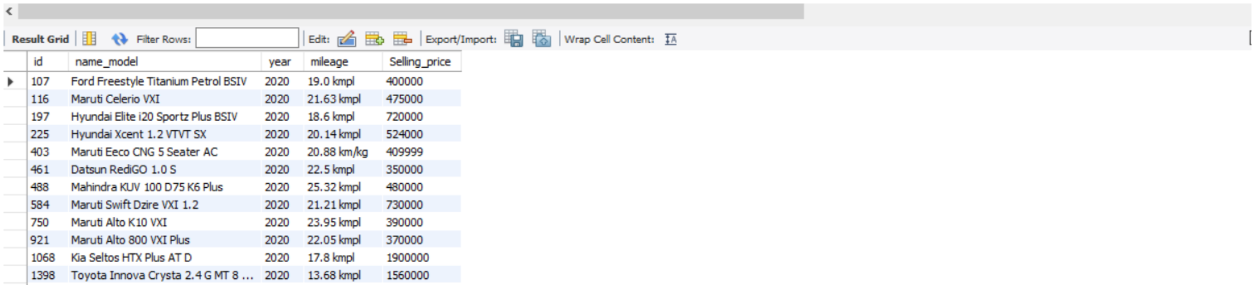
```sql
select max(year)
from used_cars
```



max(year)
2020



```sql
SELECT id, name_model, year, mileage, Selling_price
from used_cars
where year like '2020'
```



id	name_model	year	mileage	Selling_price
107	Ford Freestyle Titanium Petrol BSIV	2020	19.0 kmpl	400000
116	Maruti Celerio VXI	2020	21.63 kmpl	475000
197	Hyundai Elite i20 Sportz Plus BSIV	2020	18.6 kmpl	720000
225	Hyundai Xcent 1.2 VTVT SX	2020	20.14 kmpl	524000
403	Maruti Eeco CNG 5 Seater AC	2020	20.88 km/kg	409999
461	Datsun RediGO 1.0 S	2020	22.5 kmpl	350000
488	Mahindra KUV 100 D75 K6 Plus	2020	25.32 kmpl	480000
584	Maruti Swift Dzire VXI 1.2	2020	21.21 kmpl	730000
750	Maruti Alto K10 VXI	2020	23.95 kmpl	390000
921	Maruti Alto 800 VXI Plus	2020	22.05 kmpl	370000
1068	Kia Seltos HTX Plus AT D	2020	17.8 kmpl	1900000
1398	Toyota Innova Crysta 2.4 G MT 8 ...	2020	13.68 kmpl	1560000

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

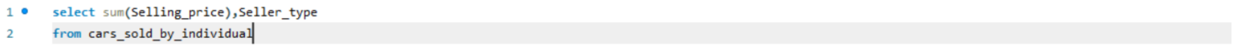
Answer:

The best car related to price and low mileage is the Maruti Omni LPG CARGO BSIII W IMMOBILISER (price: 5400 and mileage: 10.9 km/kg).

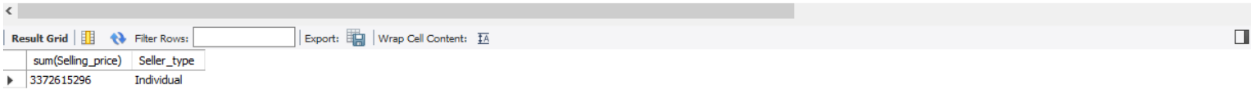
To get that result we sorted the table from low price to high price and gave every car a rank. The car with the lowest price got rank 1, the second lowest price rank 2 and so one. Then we sorted the table from low mileage to high mileage and again gave every car a rank. The car with the lowest mileage got rank 1, the second lowest mileage rank 2 and so one. We listed those ranks and then added the two ranks of every car together. The Maruti Omni LPG CARGO BSIII W IMMOBILISER got the least rank with rank 74 in price and 44 in mileage which makes a total of 118.

- **What is the total revenue from Individual sales?**

Answer: The revenue of all Individual sales is: 3372615296 ₹



```
select sum(Selling_price),Seller_type
from cars_sold_by_individual
```



sum(Selling_price)	Seller_type
3372615296	Individual