Data Science SQL Mini-Project

The objective of this project is to provide valuable business insights derived from the CARS24 dataset by writing optimized SQL queries, accompanied by the presentation of their corresponding outputs.

Dataset link- Cars24_dataset

Authored by: Samiksha Mirjha

Business Insights for cars24 company

ASSUMPTION: I have assumed that the dataset provided, is the data of the cars purchased by the company and is now in the cars24 inventory.
Which means the selling_price column in the dataset is the price at which cars24 company has purchased the cars for selling.

INSIGHT NO. 1

Reading the entire table

select * from cars24_dataset;

```
3 •   select * from cars24_dataset;
4
5
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

name	year	selling_price	km_driven	fuel	seller_type	transmission	owner	mileage	engine [CC]	max_power	seats
Hyundai i20 Asta 1.2	2007	550000	2360457	Petrol	Individual	Manual	Second Owner	18.6	1197	81.83	5
Maruti Wagon R LXI Minor	2010	194000	577414	Petrol	Individual	Manual	Second Owner	18.9	1061	67	5
Maruti Wagon R VXI BS IV	2011	229999	500000	Petrol	Individual	Manual	Second Owner	18.9	998	67.1	5
Maruti Wagon R LXI BS IV	2012	220000	360003	Petrol	Individual	Manual	Second Owner	18.9	998	67.1	5
Hyundai Sonata 2.4 GDi MT	2012	550000	330000	Petrol	Individual	Manual	Second Owner	13.44	2359	198.25	5
Hyundai Sonata 2.4 GDi MT	2012	500000	330000	Petrol	Individual	Manual	Second Owner	13.44	2359	198.25	5
Maruti Ertiga BSIV VXI	2017	700000	227000	Petrol	Individual	Manual	First Owner	17.5	1373	91.1	7
Hyundai i20 1.2 Asta	2011	220000	220000	Petrol	Individual	Manual	Fourth & Above Owner	17	1197	80	5
Maruti 800 EX	2004	70000	220000	Petrol	Individual	Manual	Second Owner	16.1	796	37	4
Honda Civic 1.8 S AT	2007	175000	218463	Petrol	Individual	Automatic	First Owner	12.9	1799	130	5
Hyundai Verna XXi ABS (Pe...	2009	340000	214000	Petrol	Individual	Manual	Second Owner	13.9	1599	103.2	5
Renault KWID RXT	2015	210000	210000	Petrol	Individual	Manual	Second Owner	25.17	799	53.3	5
Maruti Alto LX	2000	108000	206000	Petrol	Individual	Manual	Fourth & Above Owner	19.7	796	46.3	5
Hyundai i10 Magna 1.1L	2010	187000	200400	Petrol	Individual	Manual	Second Owner	19.81	1086	68.05	5
Ford Fiesta 1.4 Duratec ZXI	2008	136000	200185	Petrol	Individual	Manual	First Owner	16.6	1388	68	5
Maruti Swift Dzire 1.2 Vxi ...	2010	210000	200000	Petrol	Individual	Manual	First Owner	17.5	1197	85.8	5
Maruti Zen Estilo VXI BSIV	2010	160000	200000	Petrol	Individual	Manual	First Owner	19	998	67.1	5
Honda CR-V 2.0L 2WD AT	2006	125000	200000	Petrol	Individual	Automatic	Third Owner	13.1	1997	141.1	5
Maruti Wagon R LX	2006	65000	198000	Petrol	Individual	Manual	Second Owner	18.9	998	67.1	5
Maruti Alto LXi BSIII	2008	100000	195000	Petrol	Individual	Manual	Second Owner	19.7	796	46.3	5
Hyundai Santro Xing GLS	2008	120000	191000	Petrol	Individual	Manual	First Owner	17.92	1086	62.1	5
Hyundai Santro AT	2005	120000	190000	Petrol	Individual	Automatic	Second Owner	19.41	1458	91.5	5
Maruti Wagon R LXI	2005	70000	188000	Petrol	Individual	Manual	Fourth & Above Owner	18.9	998	67.1	5
Maruti Gypsy King Soft Top	1997	300000	186388	Petrol	Individual	Manual	Second Owner	14.8	1298	80	8

COMMENT: This is the entire dataset from cars24. Select statement returns a result set of record from the table

INSIGHT N0.2

Total numbers of cars available in the dataset/company/cars24 inventory

select count(*) as Total_cars from cars24_dataset;

```
5 •  select count(*) as Total_cars from cars24_dataset;
6
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content

Total_cars
8128

COMMENT:
The count() function is used to count the number of rows, hence we found out the total number of cars in our dataset which is 8128. We used alias(AS) here to rename the column. An alias only exists for the duration of query.

INSIGHT NO. 3

Number of cars with respect to their manufacturing year.

select year, count(*) as count_of_cars from cars24_dataset
group by year
order by count_of_cars desc;

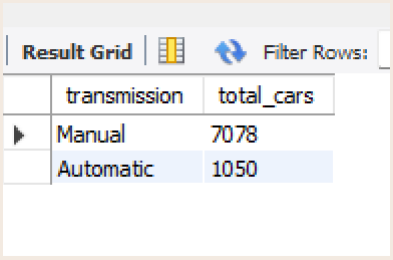
year	count_of_cars
2017	1018
2016	859
2018	807
2015	776
2013	670
2012	651
2014	621
2011	592
2019	583
2010	394
2009	246
2008	214
2007	183
2006	124
2005	97
2020	74
2004	62
2003	49
2002	27
2000	22
1999	18

COMMENT: From here we got to know that the maximum number of cars are manufactured in the year 2017.We used group by to arrange identical data into groups and order by to sort the data.

INSIGHT NO. 4

Total number of cars with respect to the transmission (Automatic and Manual).

```
select transmission, count(*) as total_cars from cars24_dataset
group by transmission
order by total_cars desc;
```


Result Grid | Filter Rows:

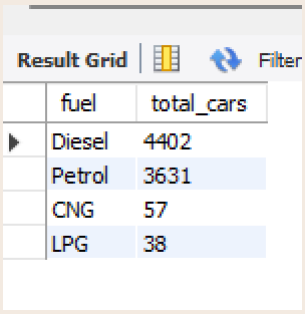
transmission	total_cars
Manual	7078
Automatic	1050

COMMENT: Here we got to know that our majority cars are manual.

INSIGHT No. 5

Total number of cars with respect to the fuel used (Petrol, Diesel,CNG,LPG).

```
select fuel, count(*) as total_cars from cars24_dataset
group by fuel
order by total_cars desc;
```


Result Grid | Filter

fuel	total_cars
Diesel	4402
Petrol	3631
CNG	57
LPG	38

COMMENT: As we can see, diesel is the most used fuel, followed by petrol .

INSIGHT NO. 6

Which type of seller (individual,dealer,or trustmark dealer) contributes the most to the inventory ?

```
select seller_type, round((count(*) * 100/(select count(*) from cars24_dataset)),2) as
Percentage_of_contribution
from cars24_dataset
group by seller_type;
```

seller_type	Percentage_of_contribution
Individual	83.24
Dealer	13.85
Trustmark Dealer	2.90

COMMENT: From here we got to know the percentage contribution of each seller. So the maximum cars are bought from the individual sellers.Subquery is used to find the total count of cars.Then dividing with the count of each seller_type cars. Round() function is used to round a number to a specified number of decimal places.

INSIGHT NO. 7

Top 10 cars which have been most frequently purchased by the company .

```
select name,count(*) as Count_of_car from cars24_dataset
group by name
order by Count_of_car desc
limit 10;
```

name	Count_of_car
Maruti Swift Dzire VDi	162
Maruti Alto 800 LXI	82
Maruti Alto LXi	80
BMW X4 M Sport X xDrive20d	62
Maruti Swift VDI	61
Maruti Swift VDI BSIV	59
Maruti Swift Dzire VXi	55
Maruti Wagon R LXI	53
Maruti Alto K10 VXI	50
Hyundai EON Era Plus	48

COMMENT: Maruti Swift Dzire VDi is the most frequently purchased car out of all with a total count of 162.We used limit in the last line of query because we are only interested for the top 10 most purchased cars.

INSIGHT NO. 8

Top 10 most affordable five seater petrol cars with mileage>20.

```
select name,selling_price from cars24_dataset
where seats = 5 and fuel = "Petrol" and mileage>20
order by selling_price
limit 10;
```

name	selling_price
Maruti Alto K10 2010-2014 VXI	110000
Hyundai EON Era	129000
Maruti Alto K10 LXI	135000
Hyundai i10 Magna	140000
Hyundai EON D Lite Plus	140000
Hyundai i10 Magna	140000
Maruti Alto K10 2010-2014 VXI	140000
Hyundai i10 Magna 1.2 iTech SE	140000
Hyundai i10 Magna	142000
Hyundai i10 Magna	142000

COMMENT: So this is the list of top 10 most affordable five seater petrol cars with mileage>20. Maruti Alto k10 2010-2014 VXI stands first out of all. Here we used multiple conditions in the 'where clause'**.**

INSIGHT NO. 9

Count of budget friendly cars in different price ranges (upto 5 lakhs) under each fuel category.

```
select fuel,count(*) as count_of_cars ,'30 thousand -80 thousand' as Price_range from
cars24_dataset
where selling_price between 30000 and 80000
group by fuel
UNION
select fuel,count(*) as count_of_cars ,'80 Thousand -1.5 lakhs' as Price_range from
cars24_dataset
where selling_price between 80000 and 150000
group by fuel
UNION
select fuel,count(*) as count_of_cars ,'1.5 lakhs-3 lakhs' as Price_range from
cars24_dataset
where selling_price between 150000 and 300000
group by fuel
UNION
select fuel,count(*) as count_of_cars ,'3 lakhs-5 lakhs' as Price_range from
cars24_dataset
where selling_price between 300000 and 500000
group by fuel
order by fuel ;
```

fuel	count_of_cars	Price_range
CNG	1	30 thousand -80 thousand
CNG	6	80 Thousand -1.5 lakhs
CNG	21	1.5 lakhs-3 lakhs
CNG	32	3 lakhs-5 lakhs
Diesel	31	30 thousand -80 thousand
Diesel	167	80 Thousand -1.5 lakhs
Diesel	782	1.5 lakhs-3 lakhs
Diesel	1150	3 lakhs-5 lakhs
LPG	3	30 thousand -80 thousand
LPG	10	80 Thousand -1.5 lakhs
LPG	25	1.5 lakhs-3 lakhs
LPG	8	3 lakhs-5 lakhs
Petrol	198	30 thousand -80 thousand
Petrol	475	80 Thousand -1.5 lakhs
Petrol	1076	1.5 lakhs-3 lakhs
Petrol	1043	3 lakhs-5 lakhs

COMMENT: We have the maximum number of cars (1150) in the price range of (3 lakhs- 5 lakhs) from the diesel type. This gives us insight regarding the number of affordable cars with respect to their fuel type.UNION operator is used to combine the data from the result of two or more SELECT command queries into a single distinct result set.

INSIGHT NO. 10

Count of cars with respect to number of seats for the following car companies(Maruti,Hyundai,Honda,Tata,Toyota)

```
select seats,count(*) as count_of_cars ,'Maruti' as car_company from cars24_dataset
where name like "%maruti%"
group by seats
UNION
select seats,count(*) as count_of_cars ,'Honda' as Price_range from cars24_dataset
where name like "%honda%"
group by seats
UNION
select seats,count(*) as count_of_cars ,'Hyundai' as Price_range from cars24_dataset
where name like "%hyndai%"
group by seats
UNION
select seats,count(*) as count_of_cars ,'Tata' as Price_range from cars24_dataset
where name like "%Tata%"
group by seats
UNION
select seats,count(*) as count_of_cars ,'Toyota' as Price_range from cars24_dataset
where name like "%Toyota%"
group by seats
order by seats;
```

seats	count_of_cars	car_company
4	95	Maruti
4	25	Tata
5	2159	Maruti
5	439	Honda
5	541	Tata
5	205	Toyota
6	3	Maruti
7	160	Maruti
7	28	Honda
7	154	Tata
7	190	Toyota
8	31	Maruti
8	1	Tata
8	92	Toyota
9	7	Tata
10	5	Tata
10	1	Toyota
14	1	Tata

COMMENT: Out of these car companies, Maruti has the highest number of five seater cars in the inventory.