1. <https://www.kaggle.com/datasets/erenakbulut/user-car-prices-barcelona-2022>
2. Clean data

Replace “Ã©” with “e”

“Ã³” with “o”

“HÃ­brido” with “Hybrid”

“Ã¡” with “a”

1. Export from Excel into SQL:

CREATE TABLE Barcelonacarsales (

[Brand] varchar (50),

[Model] varchar (100),

[Price (Euros)] int,

[Year] int,

[Mileage] int,

[Fuel] varchar(50),

[Gearbox] varchar (50),

[Location] varchar (50),

)

BULK INSERT Barcelonacarsales

FROM 'C:\Users\dustin.french\Documents\SQL Server Management Studio\Datasets\Barcelonacarsales clean.csv'

WITH (FIRSTROW = 2,

FIELDTERMINATOR = ',',

ROWTERMINATOR='\n',

BATCHSIZE=250000)

**4. Data analysis**

i. Number of unique brands, models, and locations

SELECT Count(Distinct(Brand)) [Unique Car Brands], Count(Distinct(Model)) [Unique Models], Count(Distinct(Location)) [Unique Locations]

FROM Barcelonacarsales



ii. Average Price, year, and mileage

Select avg([price (euros)]) [Average Price], AVG(Year)[Average Year], avg(mileage) [Average Mileage]

FROM Barcelonacarsales



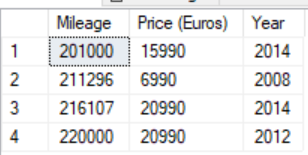
iii. The mileage, price, and year of cars that have more than 200,000 in mileage

SELECT Mileage, [Price (Euros)], Year

FROM Barcelonacarsales

WHERE Mileage > 200000

ORDER BY Mileage ASC

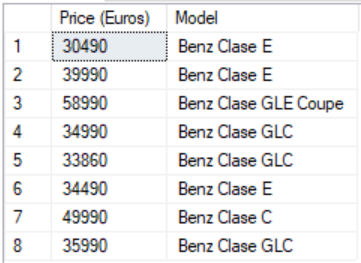


iv. The model and price of all cars with “Benz” in their model name, and price greater than 30,000 euros

SELECT [Price (Euros)], Model

FROM Barcelonacarsales

WHERE Model LIKE '%Benz%' AND [Price (Euros)] > 30000



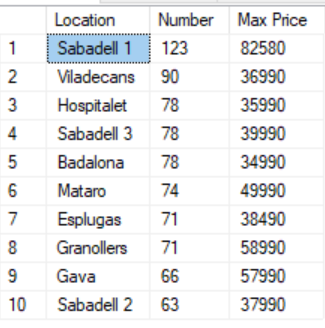
v. Location, number of cars sold in location, max price of car sold in location

SELECT Location, Count(Location) Number, MAX([Price (Euros)]) AS [Max Price]

FROM Barcelonacarsales

GROUP BY Location

ORDER BY Number Desc



vi. Top 10 cars with highest average price

SELECT TOP 10 Brand, Model, avg([price (Euros)]) AS [Avg Price]

FROM Barcelonacarsales

GROUP BY Brand, Model

ORDER BY [Avg Price] DESC



vii. The brand, model, and price of all cars sold in Sabadell 2 with an automatic transmission, categorized by price

SELECT Brand, Model, [Price (Euros)],

CASE

WHEN [Price (Euros)] < 15000 THEN 'Cheap'

WHEN [Price (Euros)] BETWEEN 15000 AND 25000 THEN 'Moderate'

ELSE 'Expensive'

END AS [Price Category]

FROM Barcelonacarsales

WHERE Location = 'Sabadell 2' AND Gearbox = 'Automatica'

ORDER BY [Price (Euros)] DESC



viii. Find average mileage of cars that have an average of over 150,000 miles.

SELECT Brand, Model, AVG(Mileage) AS [Average Mileage]

FROM Barcelonacarsales

GROUP BY Brand, Model

HAVING AVG(Mileage) > 150000

ORDER BY AVG(Mileage) DESC



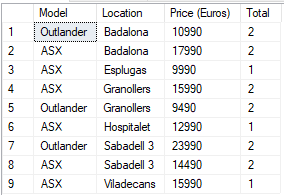
IX. Find Model, Location, and Price of Mitsubishi cars, partitioned by the number of cars sold in each location.

SELECT Model, Location, [Price (Euros)],

COUNT(Location) OVER (PARTITION by Location) AS Total

FROM Barcelonacarsales

WHERE Brand = 'Mitsubishi'



X. Find model, location, price, year, mileage of Mazda cars partitioned by average price, avg year, avg mileage of all Mazda cars sold in that location

SELECT Model, Location, [Price (Euros)], Year, Mileage,

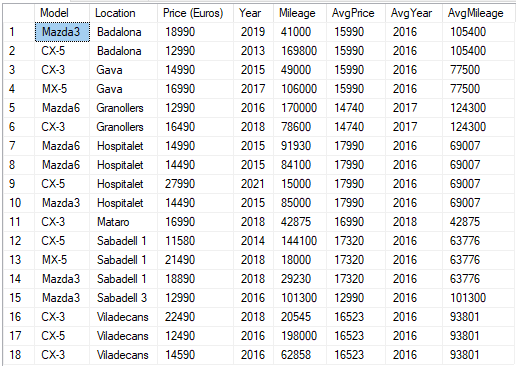
AVG([Price (Euros)]) OVER (Partition by Location) AS AvgPrice,

AVG(Year) OVER (Partition by Location) AS AvgYear,

AVG(Mileage) OVER (Partition by Location) AS AvgMileage

FROM Barcelonacarsales

WHERE Brand = 'Mazda'



Xi. Only list the Mazda cars with price greater than the average price in that location

WITH CTE\_Barcelona as

(SELECT Model, Location, [Price (Euros)], Year, Mileage,

AVG([Price (Euros)]) OVER (Partition by Location) AS AvgPrice,

AVG(Year) OVER (Partition by Location) AS AvgYear,

AVG(Mileage) OVER (Partition by Location) AS AvgMileage

FROM Barcelonacarsales

WHERE Brand = 'Mazda'

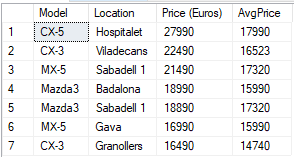
)

SELECT Model, Location, [Price (Euros)], AvgPrice

FROM CTE\_Barcelona

WHERE [Price (Euros)] > AvgPrice

ORDER BY [Price (Euros)] DESC



Xii. Only list the Mazda cars with that are newer than the average car in that location

WITH CTE\_Barcelona as

(SELECT Model, Location, [Price (Euros)], Year, Mileage,

AVG([Price (Euros)]) OVER (Partition by Location) AS AvgPrice,

AVG(Year) OVER (Partition by Location) AS AvgYear,

AVG(Mileage) OVER (Partition by Location) AS AvgMileage

FROM Barcelonacarsales

WHERE Brand = 'Mazda'

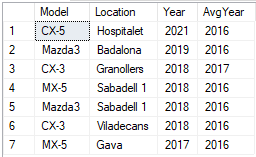
)

SELECT Model, Location, Year, AvgYear

FROM CTE\_Barcelona

WHERE Year > AvgYear

ORDER BY Year DESC



Xiii. Create procedure from vii. that shows brand, model, price, and pricing category of cars (In this case where location is Mataro and fuel is Hybrid)

ALTER Procedure [dbo].[Pricingtest]

@Location nvarchar (100),

@Fuel nvarchar (100)

AS

Create Table #temp\_pricing (

Brand varchar(50),

Model varchar(50),

[Price (Euros)] int,

[Price Category] varchar(50)

)

INSERT INTO #temp\_pricing

SELECT Brand, Model, [Price (Euros)],

CASE

WHEN [Price (Euros)] < 15000 THEN 'Cheap'

WHEN [Price (Euros)] BETWEEN 15000 AND 25000 THEN 'Moderate'

ELSE 'Expensive'

END AS [Price Category]

FROM Barcelonacarsales

WHERE Location = @Location AND Fuel = @Fuel

ORDER BY [Price (Euros)] DESC

SELECT \* FROM #temp\_pricing

ORDER BY [Price (Euros)] DESC

EXEC Pricingtest @Location = 'Mataro', @Fuel = 'Hybrid'

