**Car Sales**

This car sales data set includes information about the different cars sold by dealers in a county in the United States. The dataset has been taken from Kaggle for the purpose of predicting how many cars are sold in a month, year, or quarter and which segment of the car is considered by the customer.

Objective:

Through our dataset(car\_sales), we would draw multiple results, which might include a few scenarios to determine which age group and gender buy what type of car body or transmission. The other scenario can be to understand which dealer sells most of the cars in a quarter, month, or year.

In the car\_sales, we would consider which feature has more impact on car sales. Secondly, we have to predict car sales and check the accuracy of the prediction.

Operational Database:

ER Diagram:

Diagram

Description automatically generated

The ER Diagram consists of 6 different entitites:

Dealer which has attributes dealer\_id(Primary Key), name, age, gender, email, which consists the dealer information.

car\_sale consists of attributes sales\_unit, sale\_id(Primary Key),month, year.

car\_type entity consists of attributes type\_name, type\_id(Primary Key) which contains the body type as SUV, Sedan, Pickup, etc.

car entity consists of attributes car\_id(Primary Key), price, car\_name which contains car name and price of the car.

car\_make entity consists of attributes make, make\_id(Primary Key) which contains the information on car make brand.

Region entity contains of attributes reg\_id(Primary Key), city which informs on cars location.

Normalized table in text format:

1. Dealer {dealer\_id(PK), name, age, gender, email}
2. car\_sale{ sales\_unit, sale\_id(PK),month, year}
3. car\_type {type\_name, type\_id(PK}
4. car entity {car\_id(PK), price, car\_name}
5. car\_make {make, make\_id(PK)}
6. Region {reg\_id(PK), city}

Queries to display 1-1, 1-M, M-M relationships:

1-1

1-M

M-M

Relational Schema:

Diagram

Description automatically generated