

**TASK**

**Exploratory Data Analysis on the Automobile Data Set**

[](https://www.hyperiondev.com/)

**Introduction**

Summary of the data set

The dataset is about cars and it consists of the following information about cars:

'symboling', 'normalized-losses', 'make', 'fuel-type', 'aspiration', 'num-of-doors', 'body-style', 'drive-wheels', 'engine-location', 'wheel-base', 'length', 'width', 'height', 'curb-weight', 'engine-type', 'num-of-cylinders', 'engine-size', 'fuel-system', 'bore', 'stroke', 'compression-ratio', 'horsepower', 'peak-rpm', 'city-mpg', 'highway-mpg', 'price'.

**DATA CLEANING**

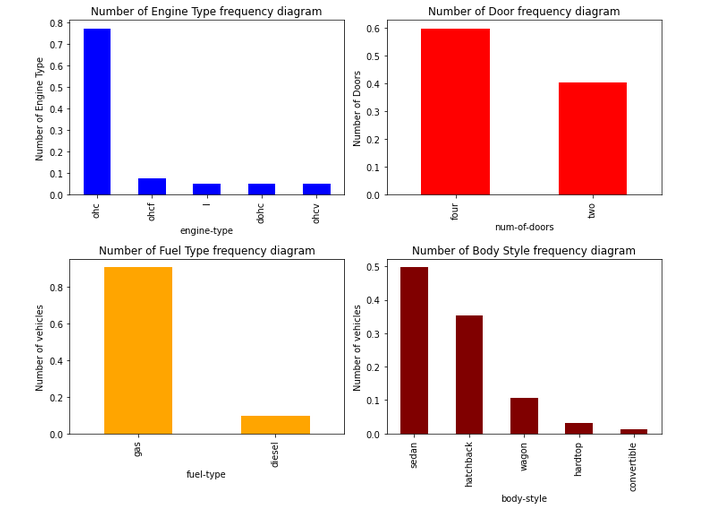
# SUMMARY OF THE METHODS AND VISUALISATIONS DONE DURING DATA CLEANING

The data cleaning process includes removing any duplicates, removing rows with missing data, changing data types, and data wrangling to make sure the car makes have unique names.

**MISSING DATA**

The missing data were in a form of question marks and they were replaced with numpy NaN and then deleted using the *dropna()* function.

**DATA STORIES AND VISUALISATIONS**

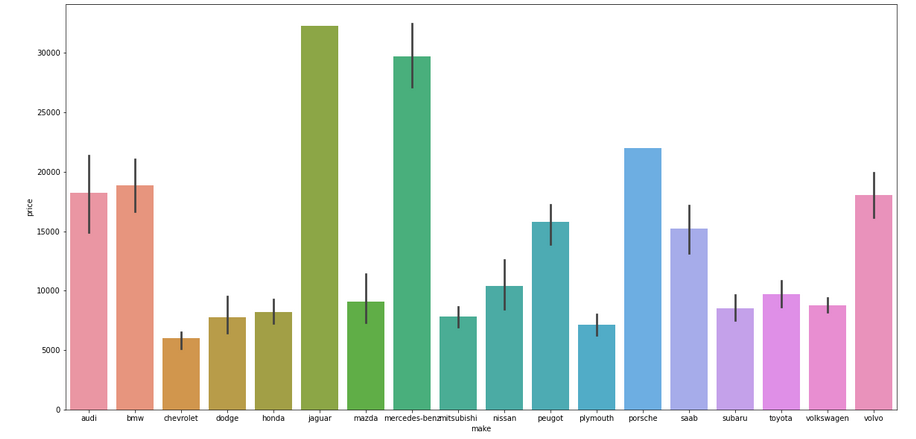


Graph 1: More than 70 % of the vehicle has Ohc type of Engine

Graph 2: 57% of the cars have 4 doors and 43% have 2 doors

Graph 3: Gas is preferred by 85 % of the vehicles

Graph 4: Most produced vehicle are of body style sedan around 48%, followed by hatchbacks with 32%.

****

**Mercedez-Benz, Jaguar, and Porshe produce the most expensive cars more than 20000 USD.**

**Chevrolet, Dodge, Honda, Mitsubishi, Nissan, Plymouth Subaru, and Toyota produce budget models with lower prices most of the car companies produce cars below 20000.**

**A graph showing different colored squares

Description automatically generated**

**Convertible car models are the most expensive in prices, followed by sedan and hardtop.**

**A graph of gas prices

Description automatically generated**

**It makes sense that most of the cars in the dataset use gasoline, looking at the fact that most of the produced cars are small cars (sedans, hatchbacks, convertibles and etc).**

**THIS REPORT WAS WRITTEN BY : MORONGWA THOBEJANE**

