

**TASK**

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

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

**Introduction**

The Dataset is a collection of data about cars, this includes the company manufacturers to the design and specifications of each car.

**DATA CLEANING**

* Removing columns that are not needed for processing by dropping them.
* Removing duplicate rows

**MISSING & INCORRECT FORMATED DATA**

* Removing rows with missing data, this included rows where the value = ‘?’
* Some of the columns needed to be converted to integers as well which I converted to int64

**DATA STORIES AND VISUALISATIONS**

**most expensive cars:** Mercedes-benz, BMW and Porsche

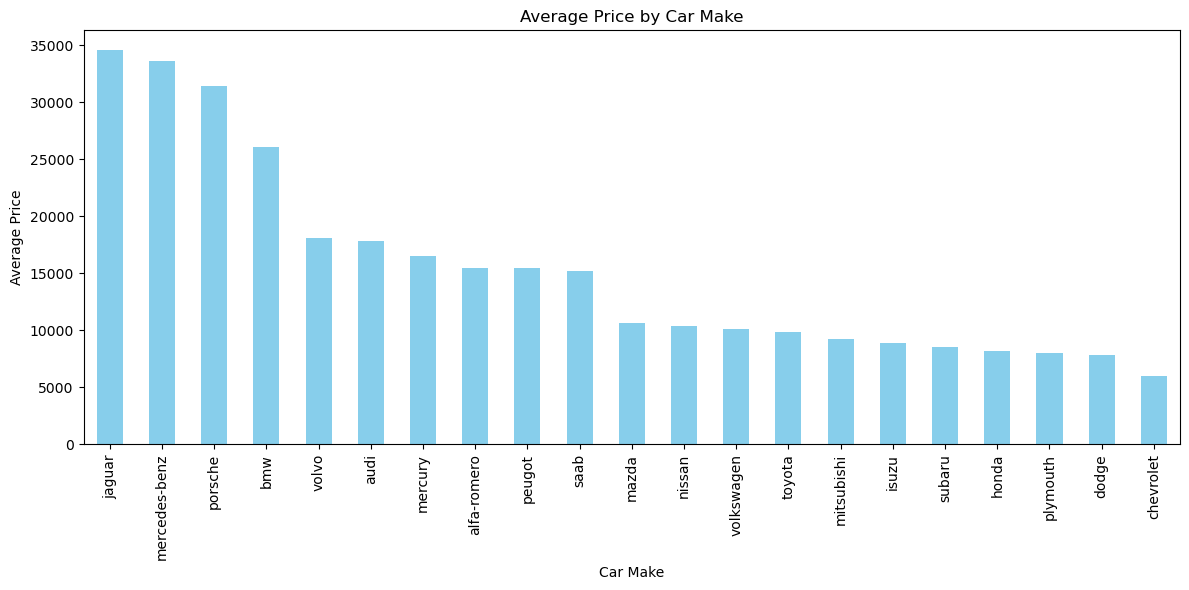
**Fastest cars:** Jaguar, Porsche, Nissan

**Lowest horsepower cars**: Chevrolet, Volkswagen, Nissan, Toyota

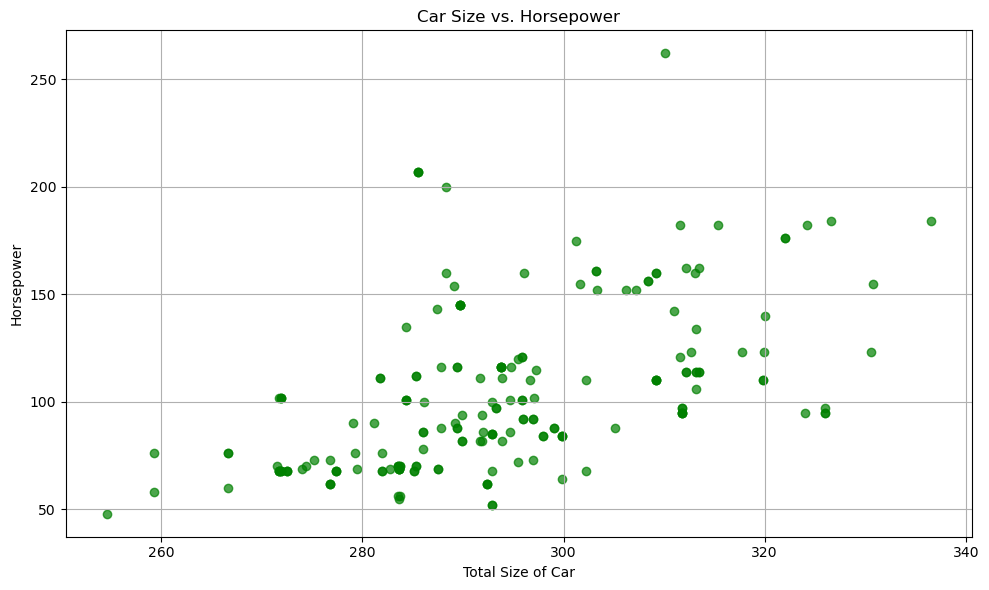
**Lowest insurance risk cars:** alfa-romero, dodge, mazda

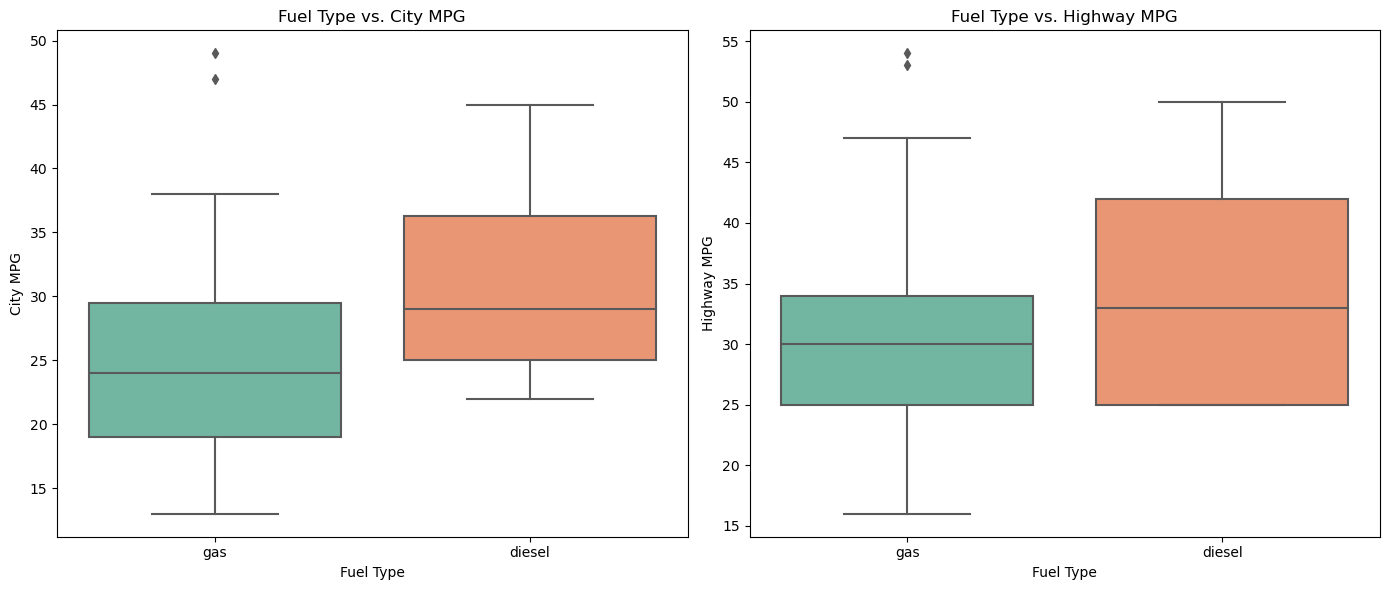
**Highest insurance risk cars:** volvo, dodge, Mercedes\_benz

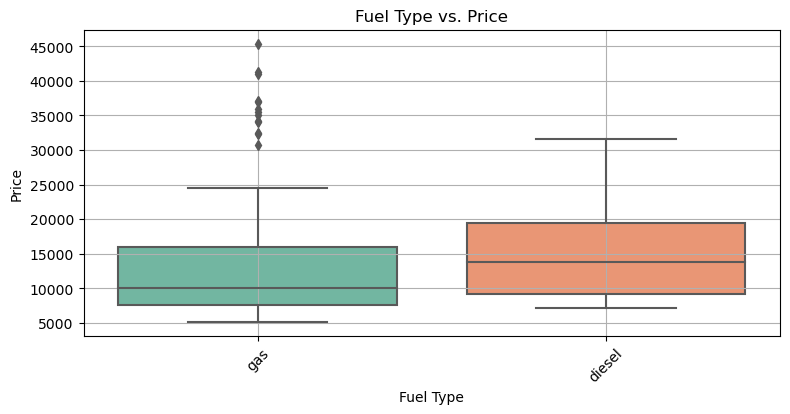
**Visualisations:**

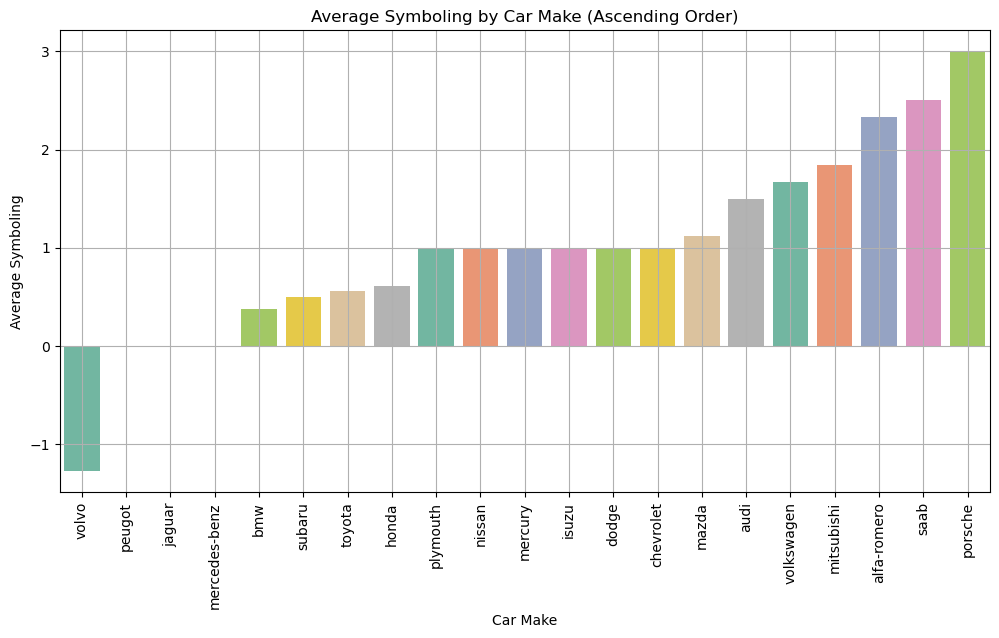


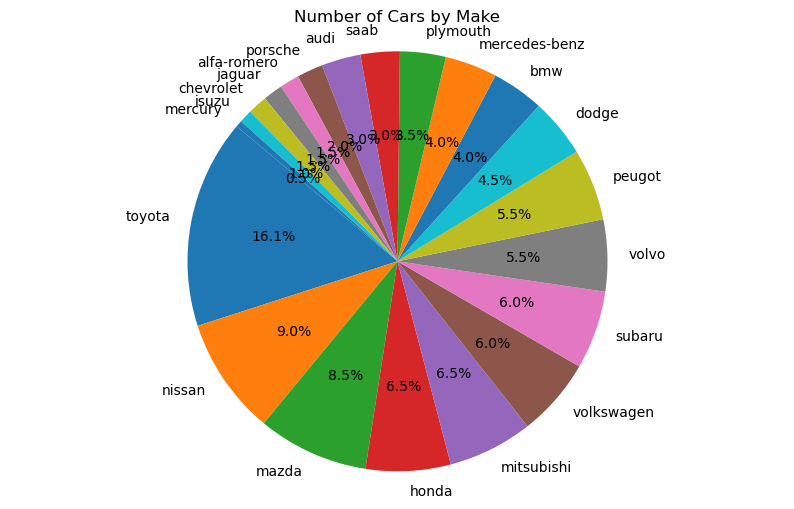
**Jaguar shows highest average price and Chevrolet shows cheapest average price.**

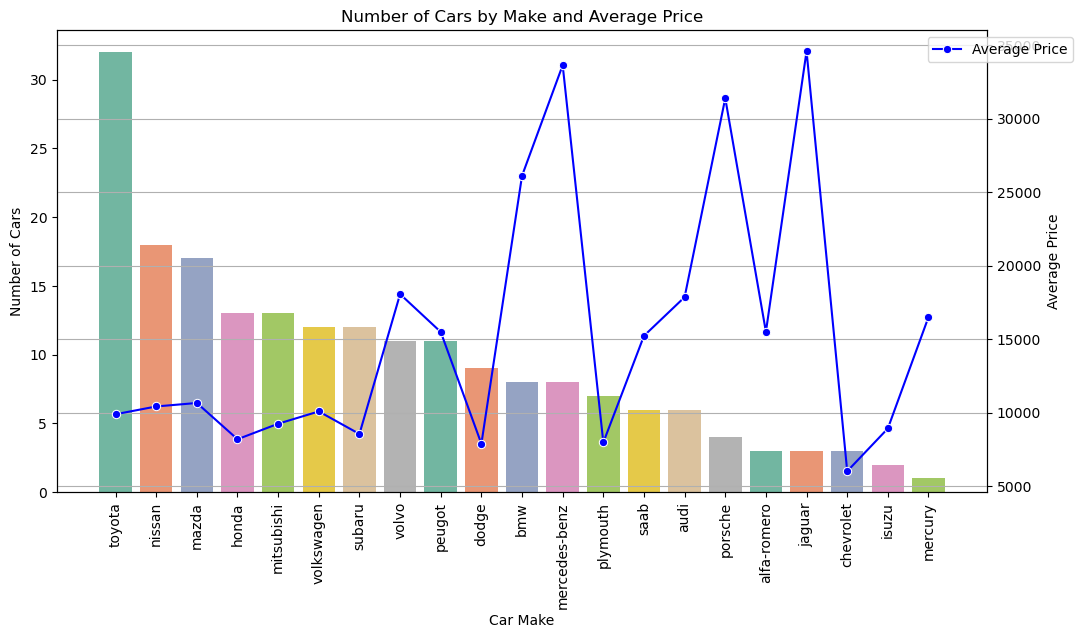
**The graph shows horsepower to car size, there is a trend that the larger the car the more horsepower is required**

**The graph shows that diesel has a higher city & highway MPG than Gas**

**The graph shows that diesel cars are more expensive however the previous graph shows that the mpg is higher for diesel thus over time it may be a better option gas prices increase**

**The graph shows which car makes have low to high insurance risk rating (Symbol), Volvo has a -1 rating and Porsche has a 3**

**Toyota has a higher variety of cars than all other car makes, The insurance risk for Toyota as seen in an above graph 'Average Symbolling by Car Make' is below average maybe because of the large variety of cars and less focused designs or possibly that the variety of makes dominate most of the market meaning more cars have a chance of accidents/ theft.**

**The graph shows that some of the manufacturing companies that make less amount of cars sell them for a lot more, even though dominating the market with variety may generate revenue, concentrating on fewer but more professional cars may generate higher profits**

**THIS REPORT WAS WRITTEN BY: RYAN COX**

