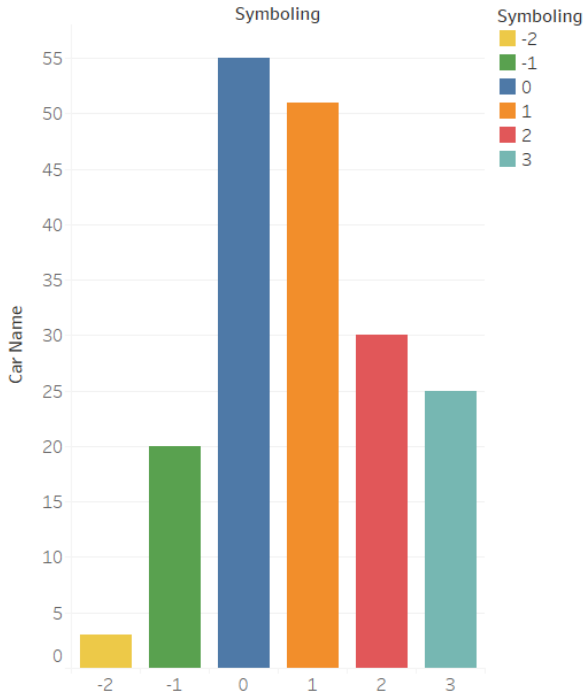
DATA DICTONARY		
1	**Car_ID**	Unique id of each observation (Interger)
2	**Symboling**	Its assigned insurance risk rating, A value of +3 indicates that the auto is risky, -3 that it is probably pretty safe.(Categorical)
3	**carCompany**	Name of car company (Categorical)
4	**fueltype**	Car fuel type i.e gas or diesel (Categorical)
5	**aspiration**	Aspiration used in a car (Categorical)
6	**doornumber**	Number of doors in a car (Categorical)
7	**carbody**	body of car (Categorical)
8	**drivewheel**	type of drive wheel (Categorical)
9	**enginelocation**	Location of car engine (Categorical)
10	**wheelbase**	Weelbase of car (Numeric)
11	**carlength**	Length of car (Numeric)
12	**carwidth**	Width of car (Numeric)
13	**carheight**	height of car (Numeric)
14	**curbweight**	The weight of a car without occupants or baggage. (Numeric)
15	**enginetype**	Type of engine. (Categorical)
16	**cylindernumber**	cylinder placed in the car (Categorical)
17	**enginesize**	Size of car (Numeric)
18	**fuelsystem**	Fuel system of car (Categorical)
19	**boreratio**	Boreratio of car (Numeric)
20	**stroke**	Stroke or volume inside the engine (Numeric)
21	**compressionratio**	compression ratio of car (Numeric)
22	**horsepower**	Horsepower (Numeric)
23	**peakrpm**	car peak rpm (Numeric)
24	**citympg**	Mileage in city (Numeric)
25	**highwaympg**	Mileage on highway (Numeric)
26	**price(Dependent variable**	Price of car (Numeric)

Car Data set Visualizations are as follows:-

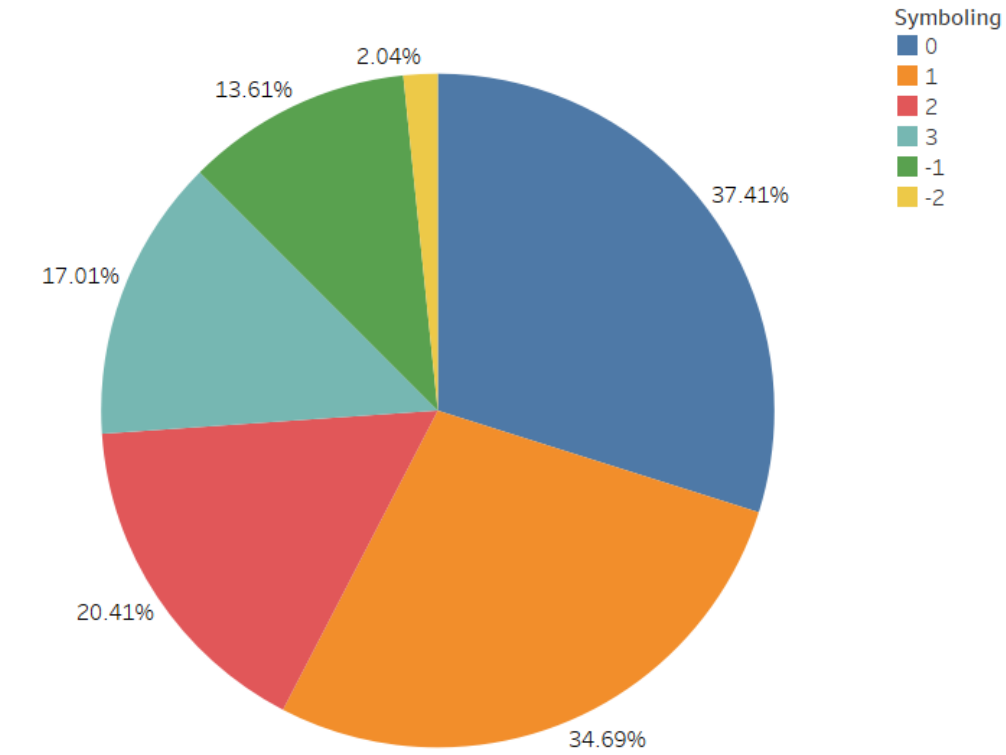
Let's see the count of automobile in each category and percent share of each category.

symboling Vs Count of cars



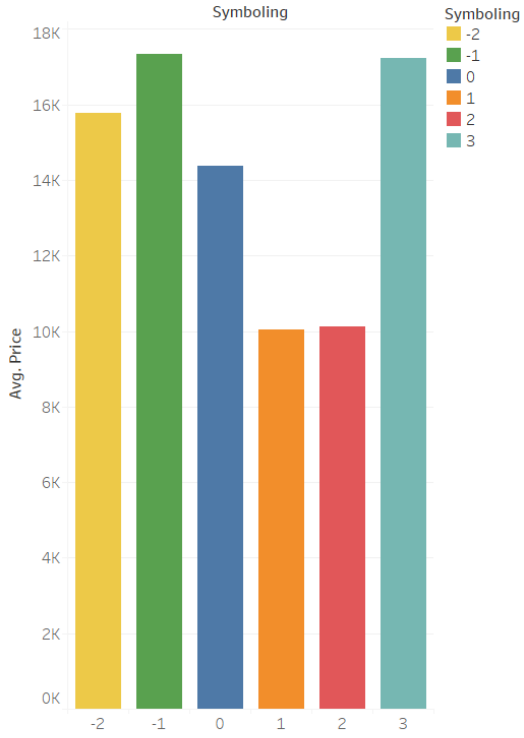
Distinct count of Car Name for each Symboling. Color shows details about Symboling.

symboling : Its assigned insurance risk rating
A value of +3 indicates that the auto is risky,
-3 that it is probably pretty safe.(Categorical)



Symboling (color) and % of Total Distinct count of Car Name (size).

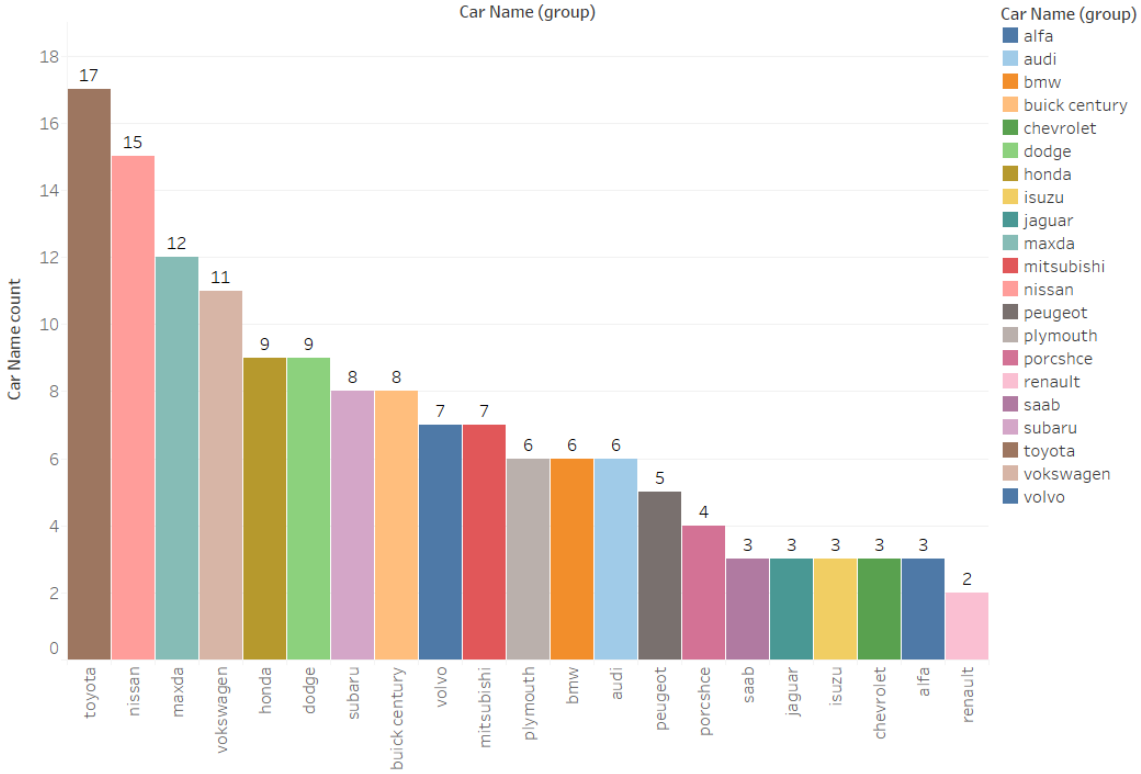
Let's see average price of cars in each symbol category.



Average of Price for each Symboling. Color shows details about Symboling.

INTERPRETATIONS
More than 50% of cars are with symbol 0 or 1.
Average price of car is lower for 0,1 & 2 symbol category.

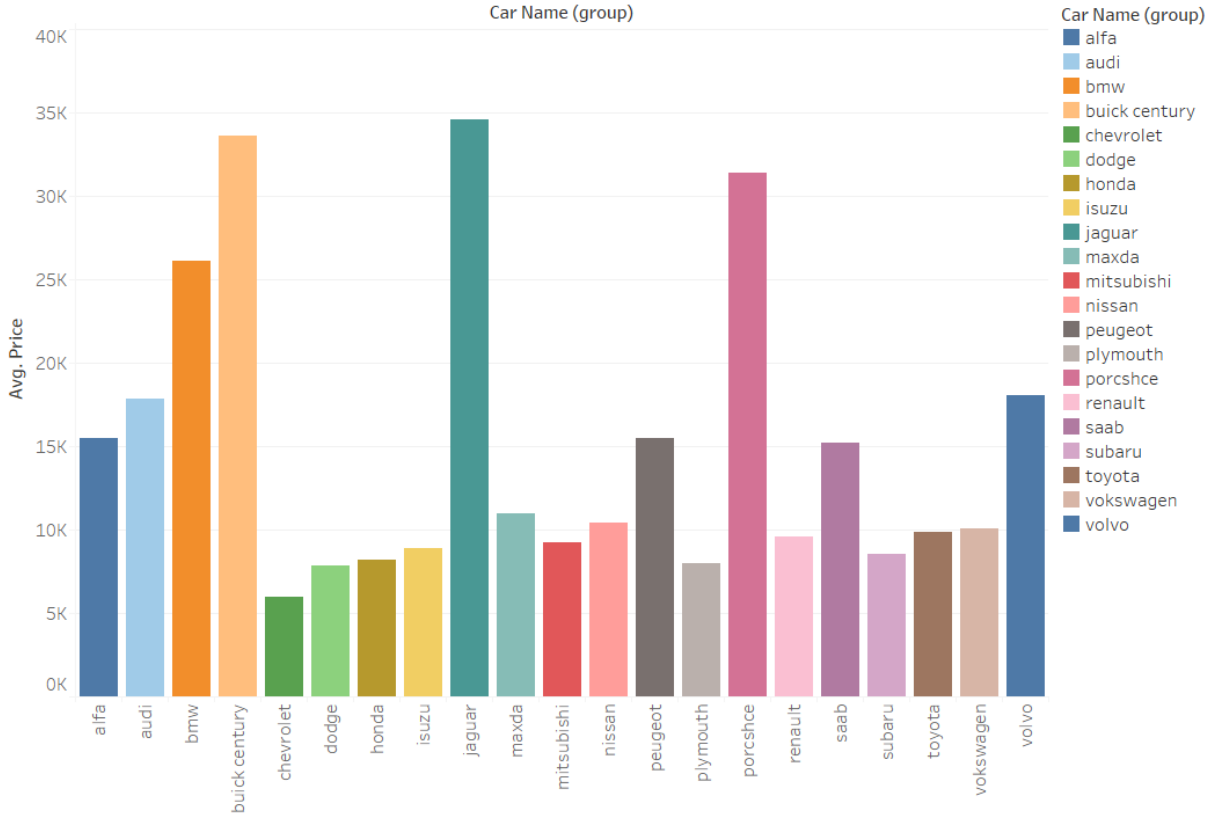
Let's see companies and their no of models.



Distinct count of Car Name count for each Car Name (group). Color shows details about Car Name (group).

Inference
Toyota, a Japanese company has the most no of models.

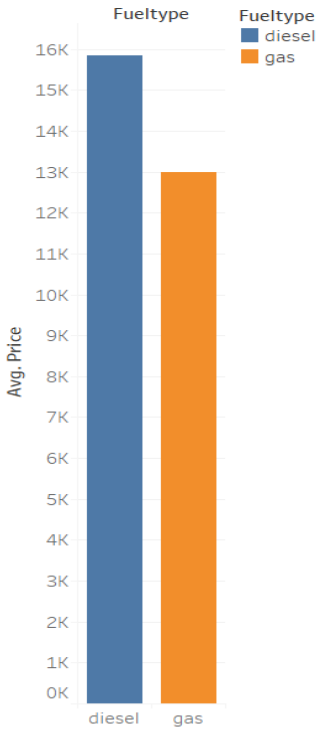
Let's see average car price of each company.



Average of Price for each Car Name (group). Color shows details about Car Name (group).

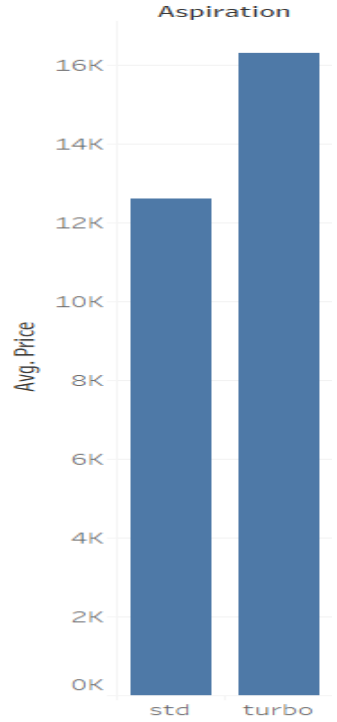
Interpretations
Toyota has considerably high no of models in the market.
Brands can be categorized as Luxury, Mid Ranged, Budget based on their average price.
Avg_price < 10000 = budget
10000 < Avg_price < 20000 = Mid_Range
Else 20000 > Avg_price = Luxury

Let's see how price varies with Fuel Type



Average of Price for each Fueltype. Color shows details about Fueltype.

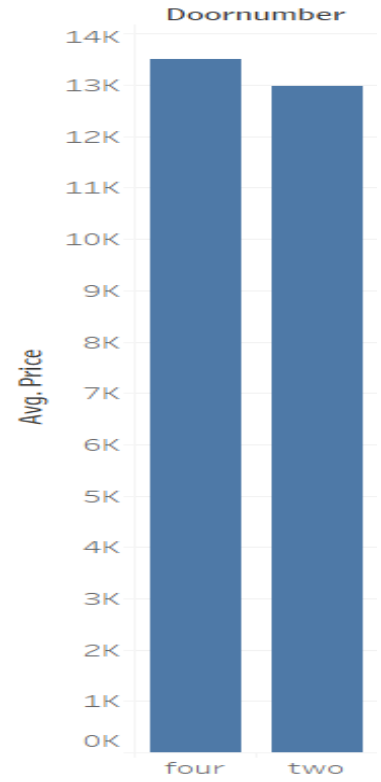
Inference
Diesel cars are priced more than gas cars.



Average of Price for each Aspiration.

Inference
Cars with turbo aspiration engine are priced more than standard ones.

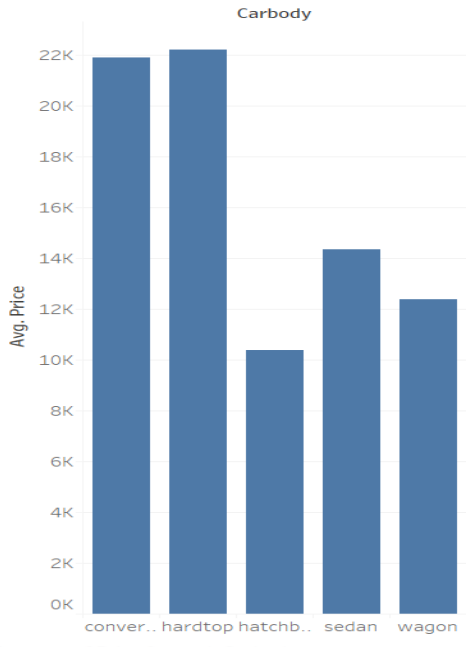
Door Numbers



Average of Price for each Doornumber.

Inference
Number of doors doesn't seem to have much effect on price.

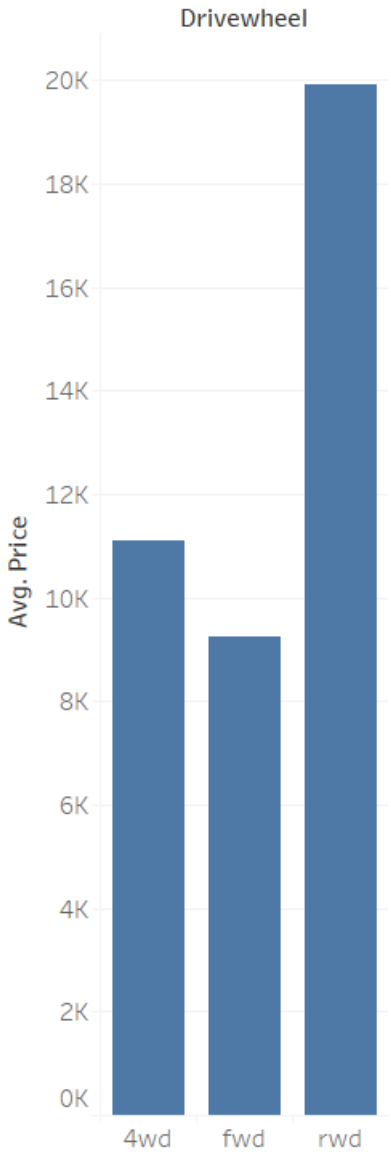
Car Body



Average of Price for each Carbody.

Inference
Hardtop and convertible are the most expensive whereas hatchbacks are the cheapest.

Drivewheel

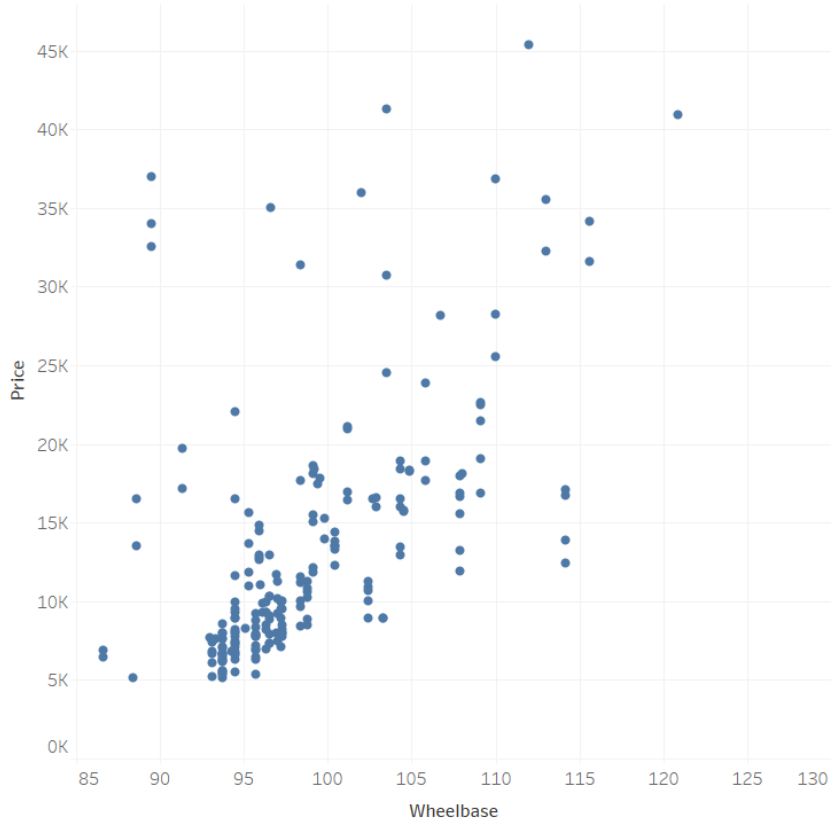


Average of Price for each Drivewheel.

Inference

Cars with Rear wheel drive have a higher price value.

Wheel base

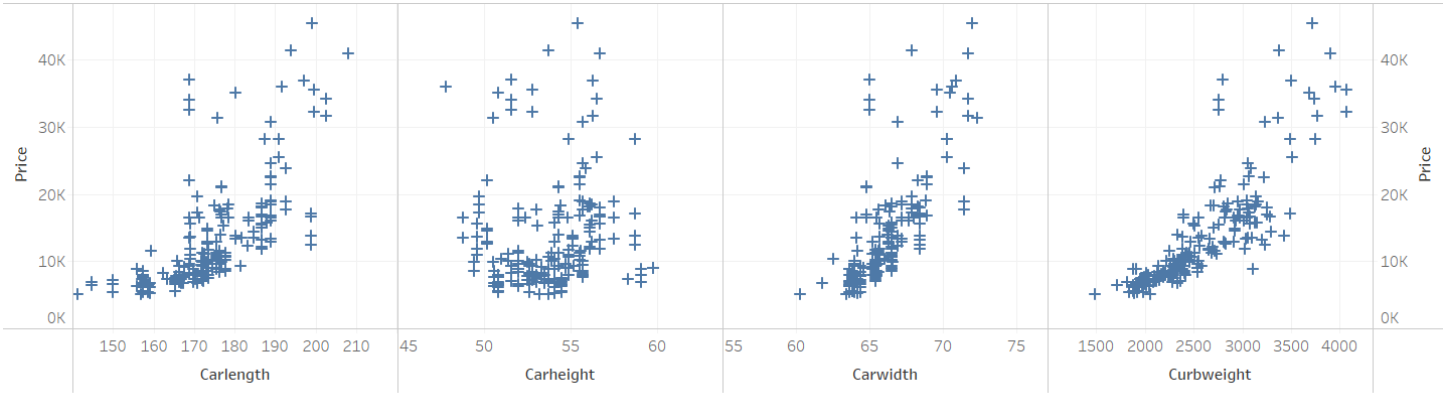


Wheelbase vs. Price.

```
Most cars has a wheel base around 95 inches.
Price has a slight positive correlation with wheelbase
```

Car Dimensions

Let's see how price varies with Car's length, width,height and weight.



Carlength, Carheight, Carwidth and Curbweight vs. Price and Price.

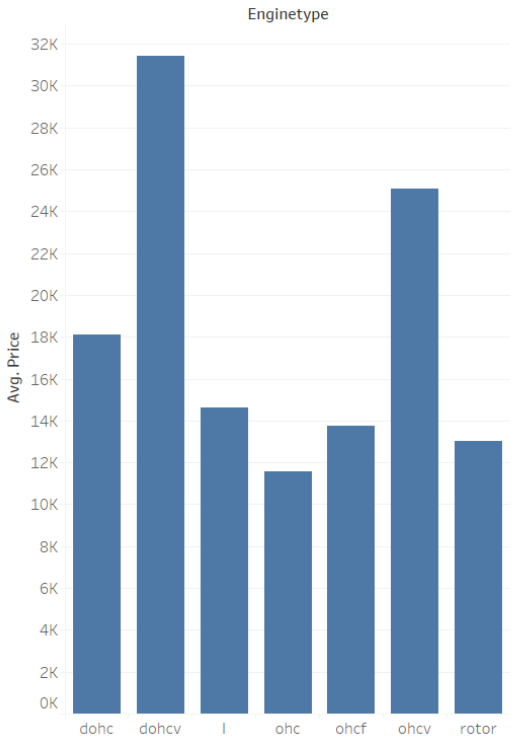
Inference

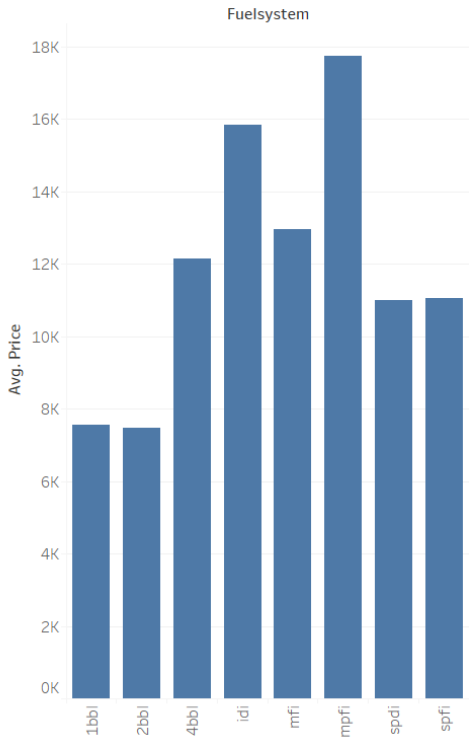
Length width and weight of the car is positively related with the price.

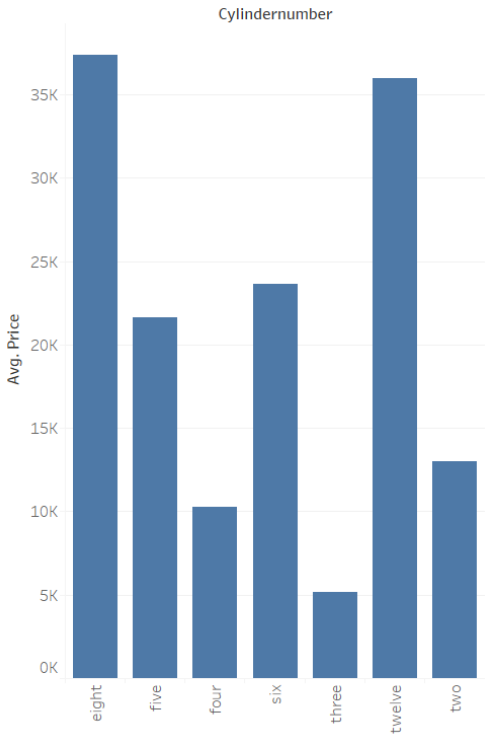
There is not much of a correlation with Height of the car with price.

Engine Specifications

Engine Type, Cylinder, Fuel System







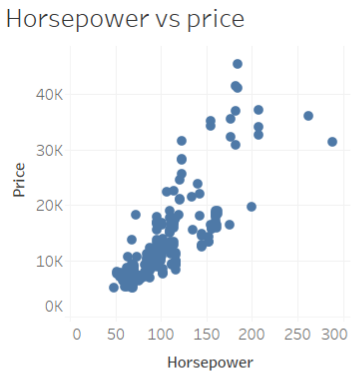
Inference

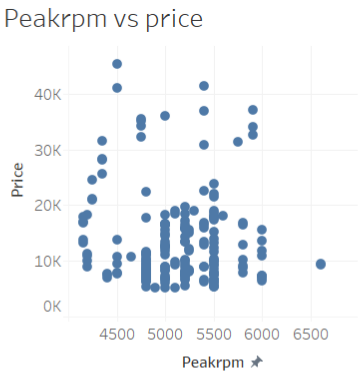
DOHCV and OHCV engine types are priced high.

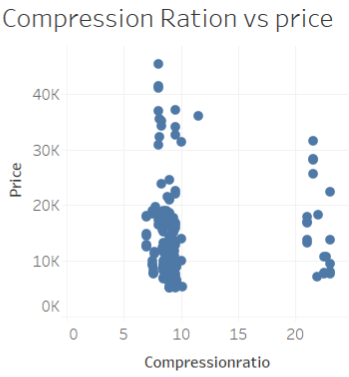
Eight and twelve cylinder cars have higher price.

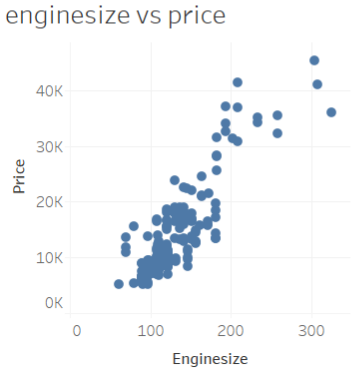
IDI and MPFI fuel system have higher price.

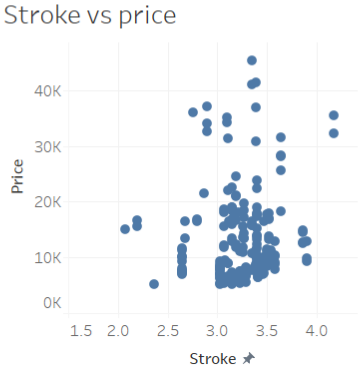
Engine Size, Bore Ratio, Stroke, Horsepower & Compression Ratio

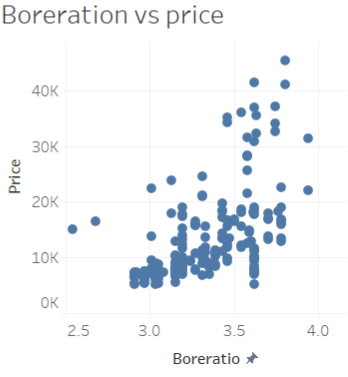








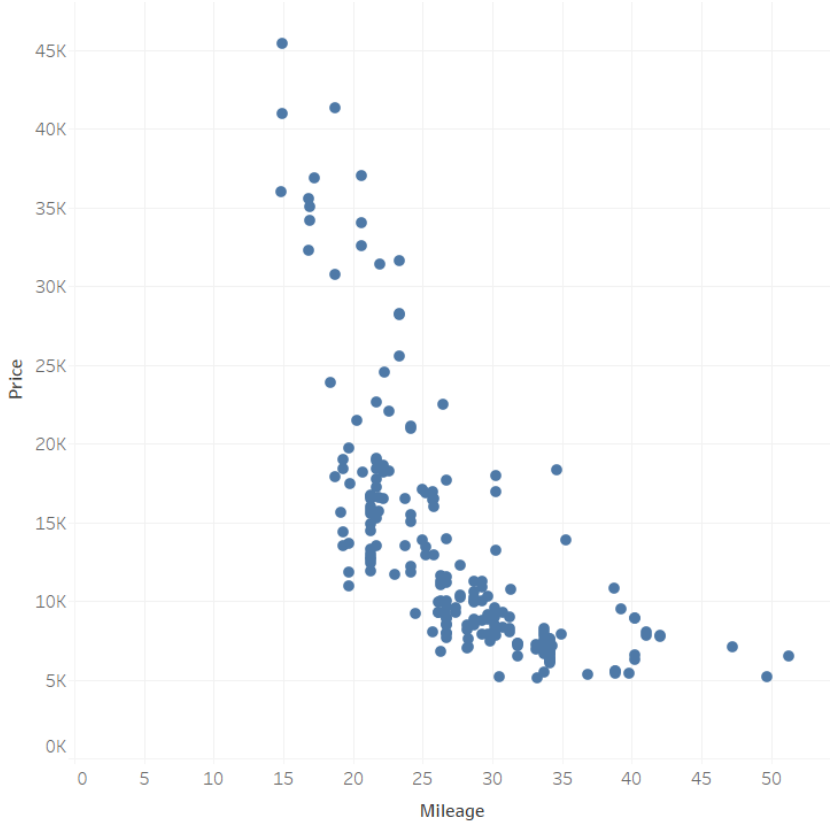




Size of Engine, bore ratio, and Horsepower has positive correlation with price.

City Mileage & Highway Mileage

A single variable mileage can be calculated taking the weighted average of 55% city and 45% highways.
Let's see how price varies with mileage.



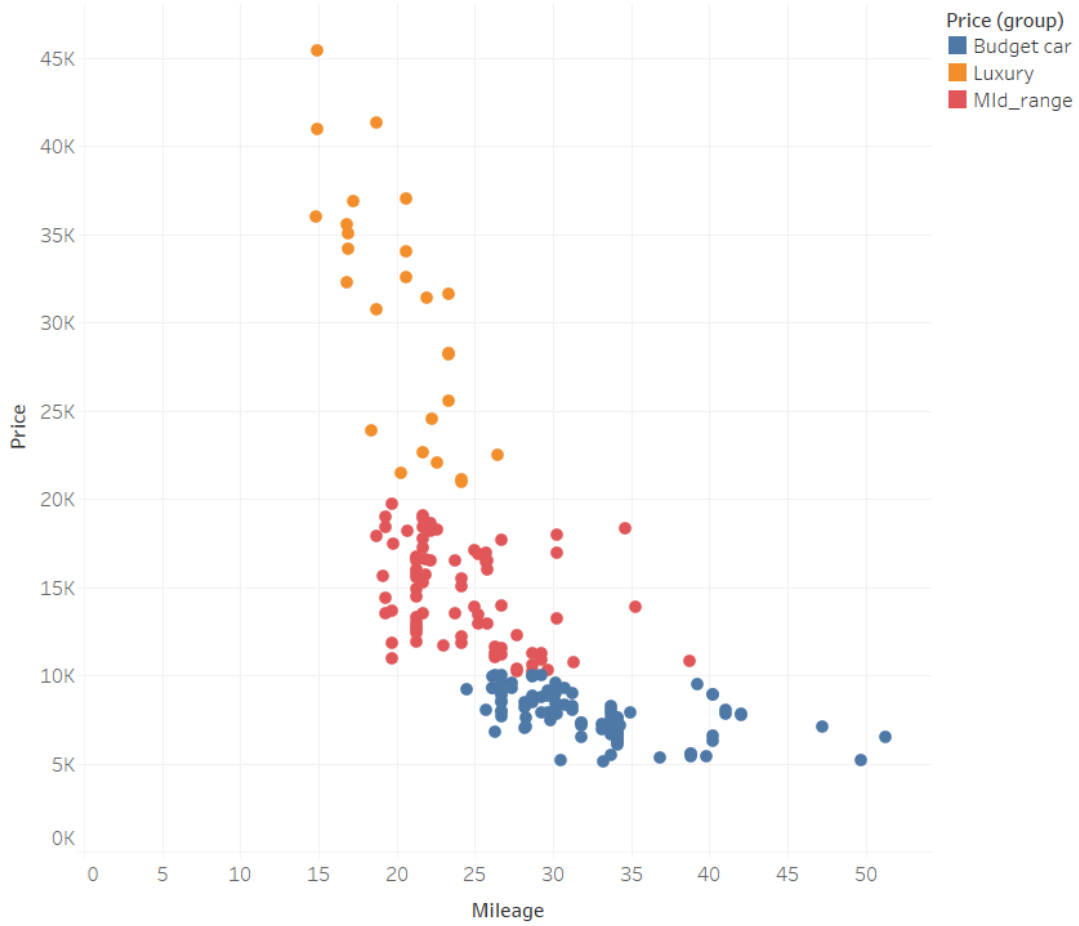
Mileage vs. Price.

Inference
Mileage has a negative correlation with price.

Bivariate Analysis

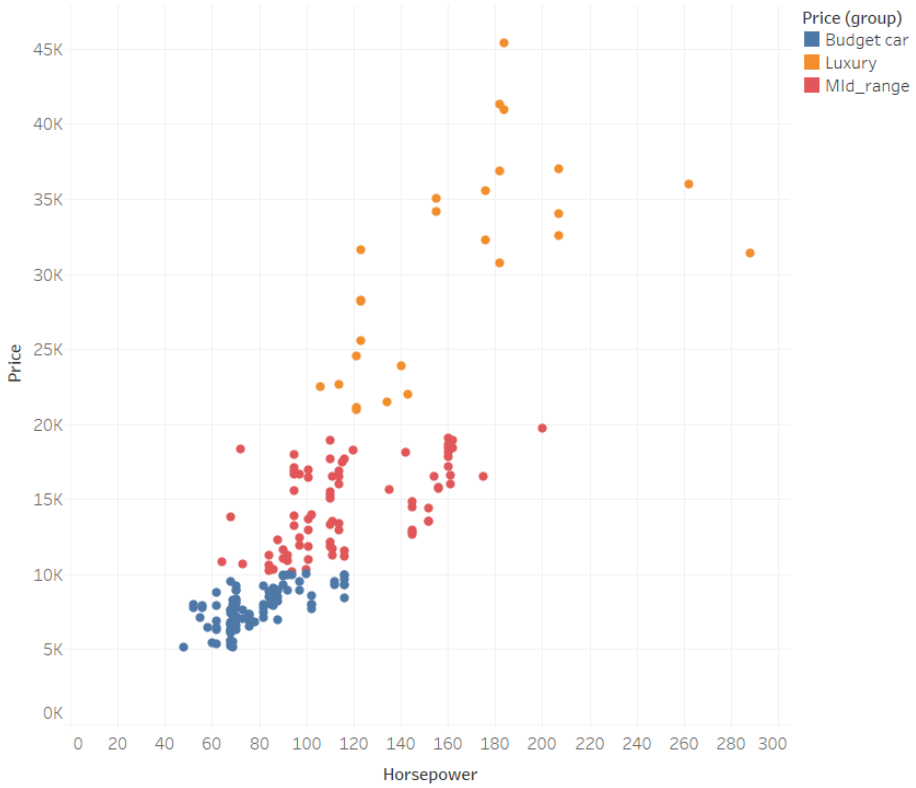
Brand Category – Mileage

t is expected that luxury brands don't care about mileage. Let's find out how price varies with brand category and mileage.



Mileage vs. Price. Color shows details about Price (group).

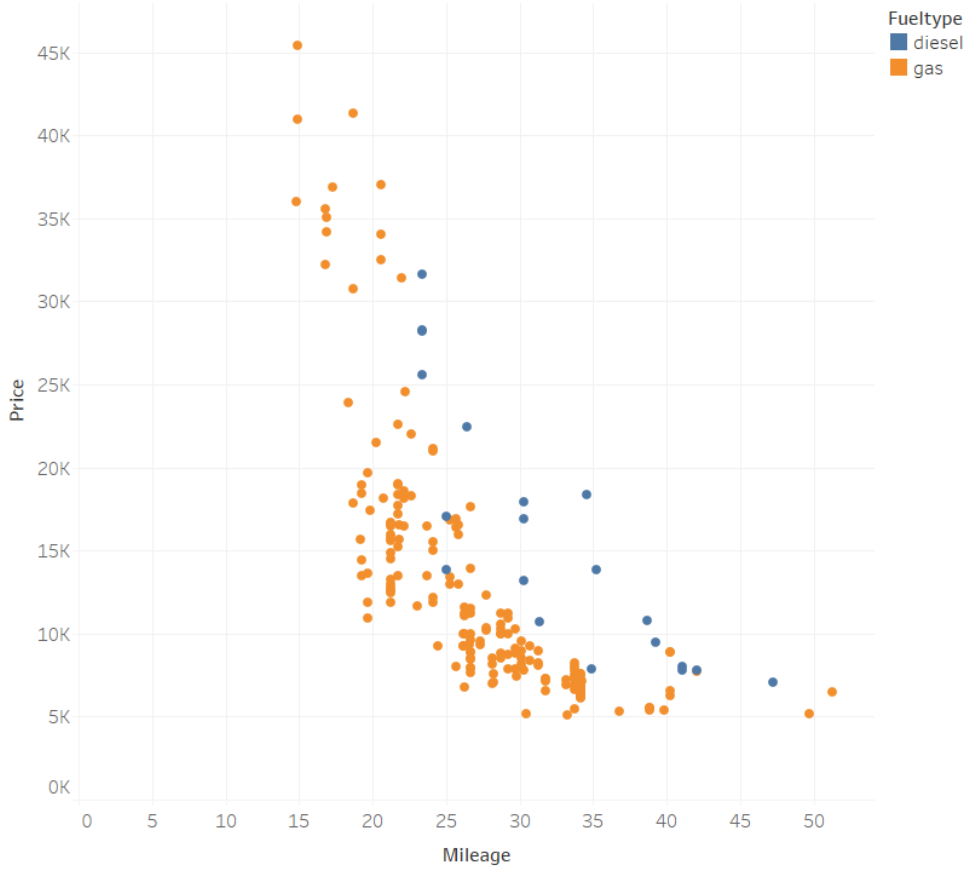
Brand Category - Horsepower



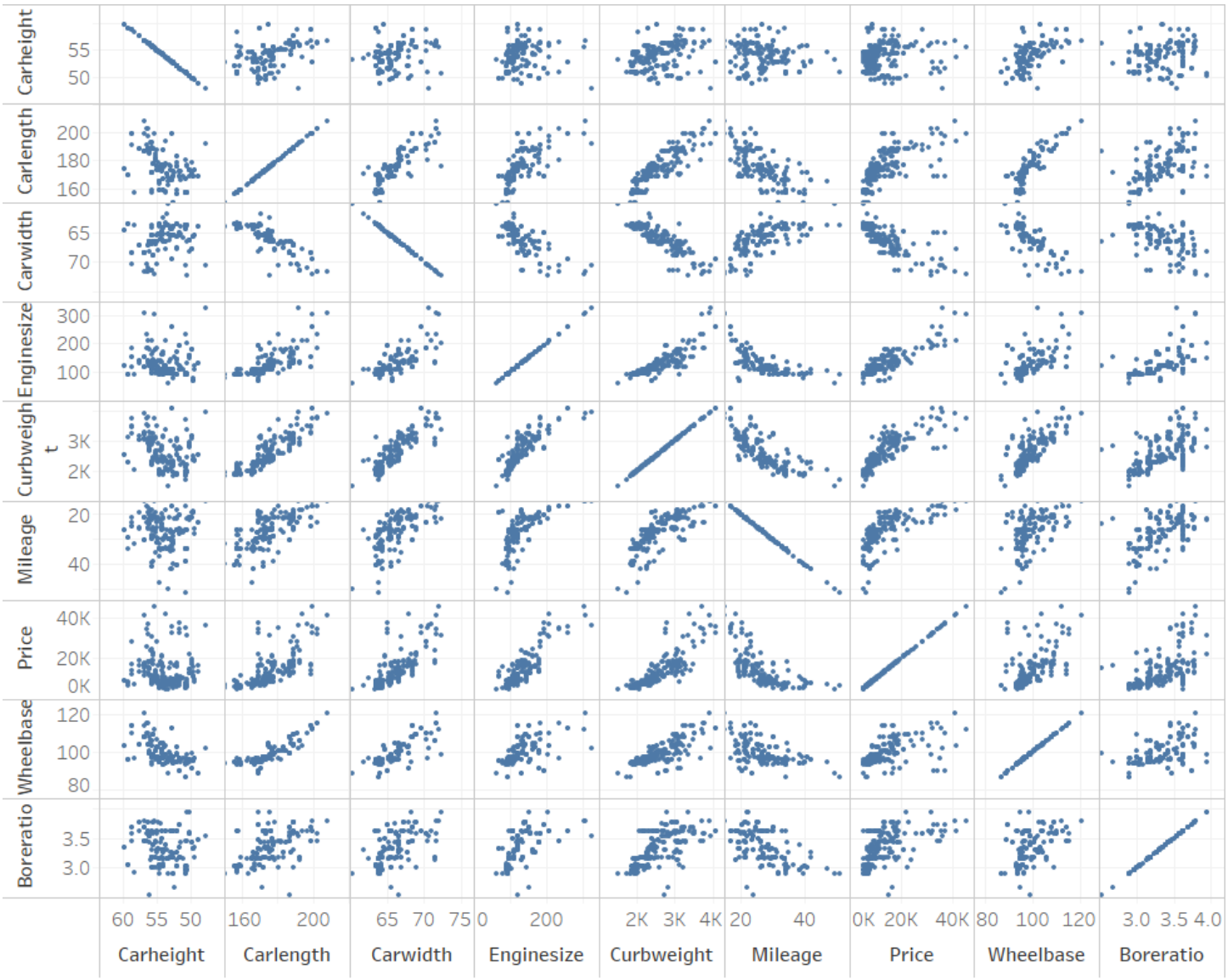
Horsepower vs. Price. Color shows details about Price (group).

Luxury cars have higher horsepower and less mileage

Mileage - Fuel Type



Mileage vs. Price. Color shows details about Fueltype.



Carheight, Carlength, Carwidth, Enginesize, Curbweight, Mileage, Price, Wheelbase and Boreratio vs. Carheight, Carlength, Carwidth, Enginesize, Curbweight, Mileage, Price, Wheelbase and Boreratio.