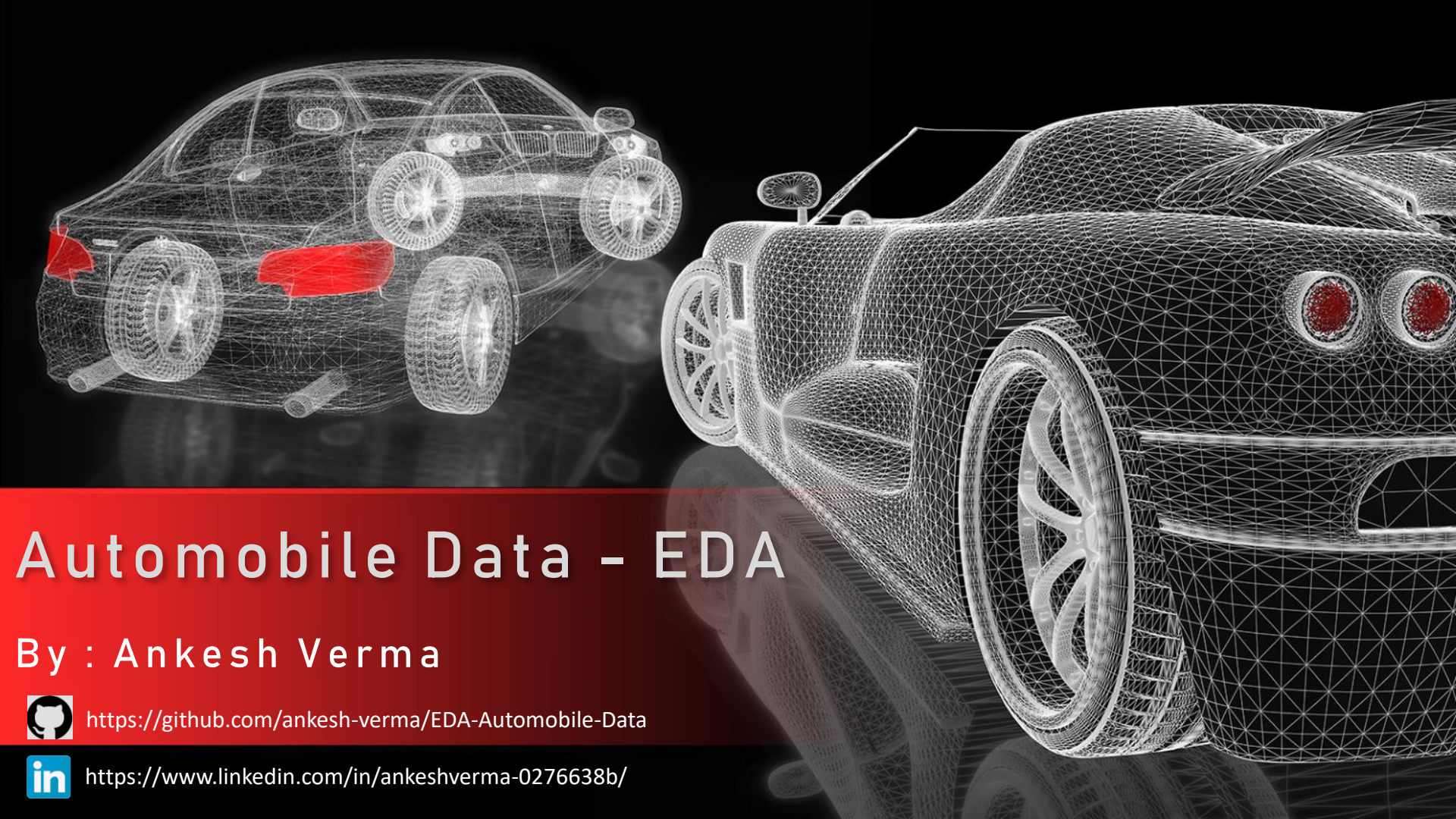


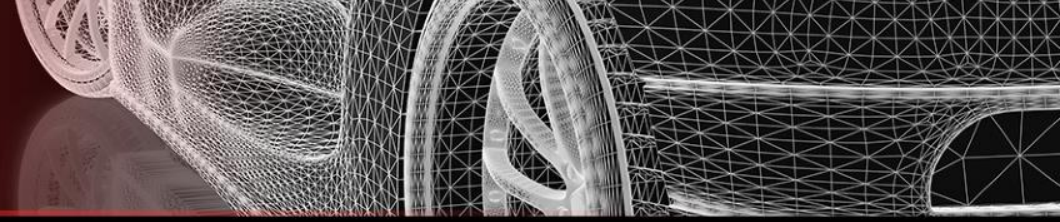
Automobile Data - EDA

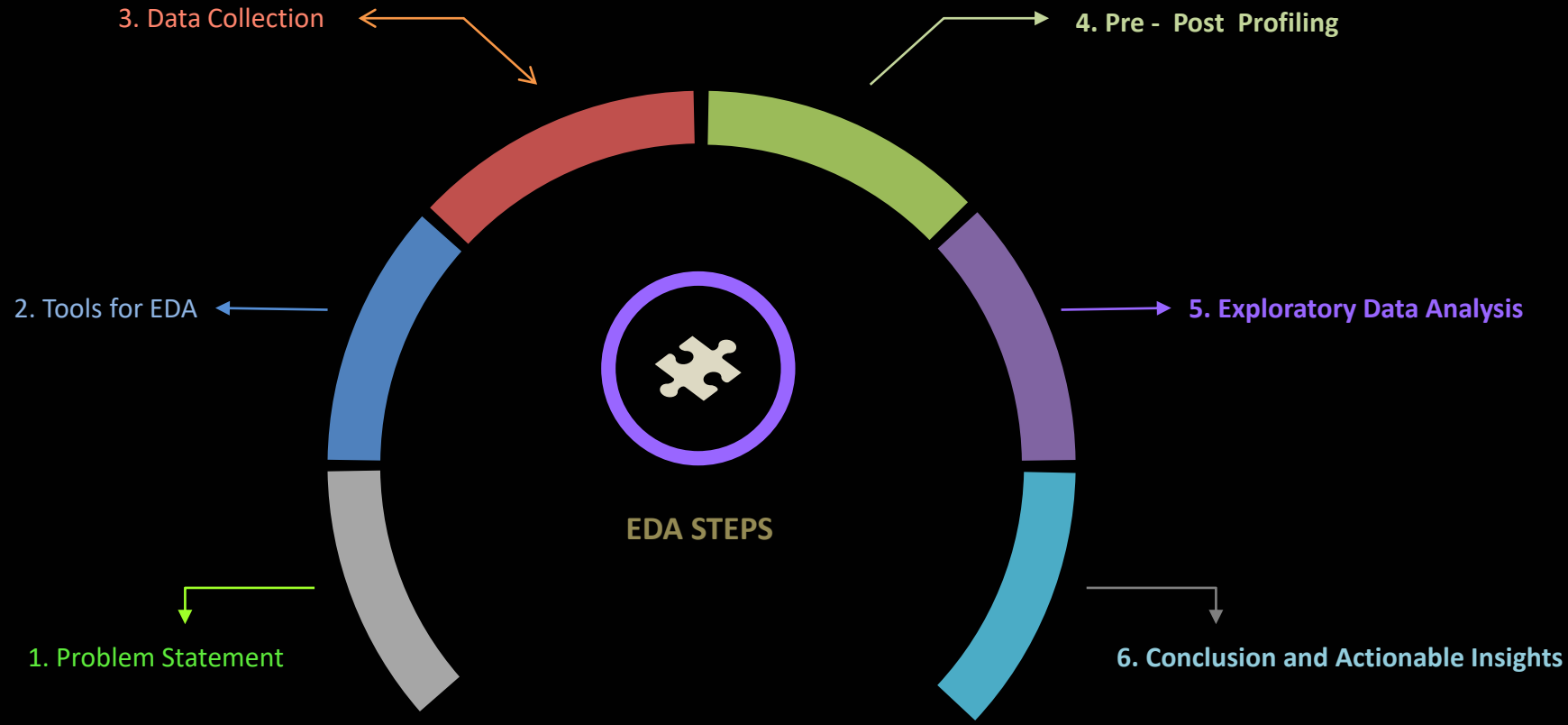
By : Ankesh Verma

https://github.com/ankesh-verma/EDA-Automobile-Data

https://www.linkedin.com/in/ankeshverma-0276638b/

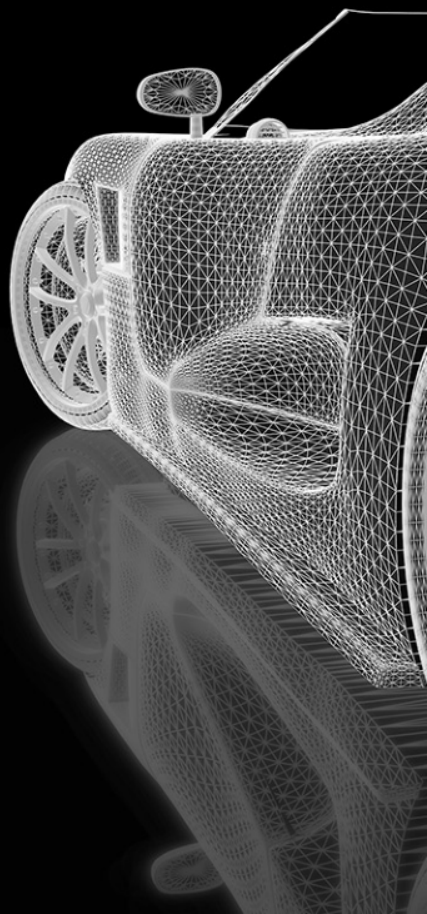
EDA STEPS



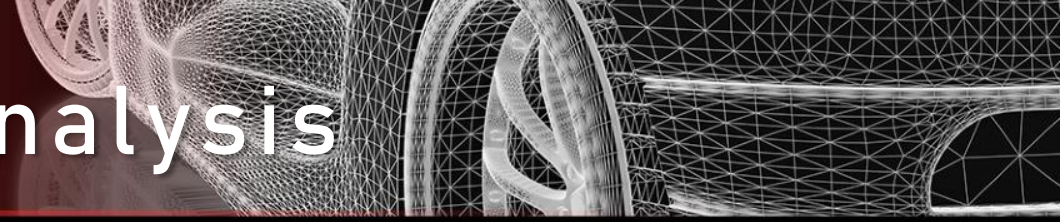

3. Data Collection
4. Pre - Post Profiling
2. Tools for EDA
5. Exploratory Data Analysis
EDA STEPS
1. Problem Statement
6. Conclusion and Actionable Insights

Problem Statement :

- Do various Configuration of cars like Horsepower, curb-weight, Body-style etc. helps in determining car prices ?



Tools Used For Analysis



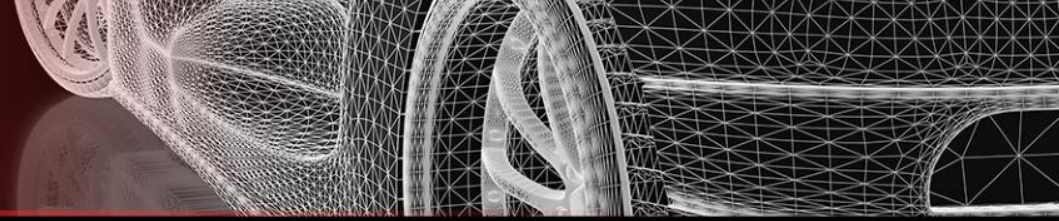
MANIPULATION

- Pandas
- Numpy

VISUALIZATION

- Matplotlib
- Seaborn

About the Data



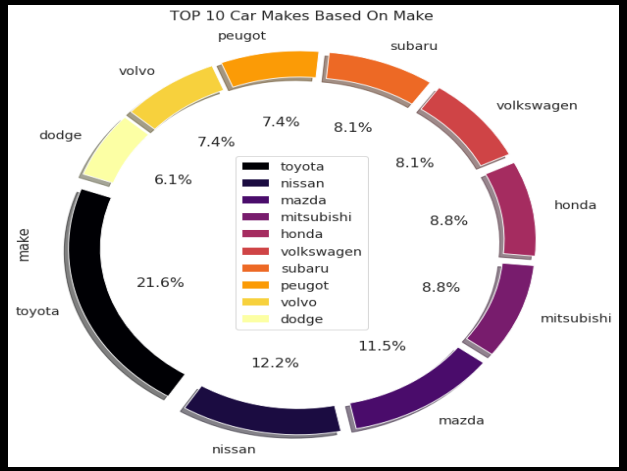
Contents: Cars specification like body – style , aspiration ,horsepower , engine and price.

Data Volume: 205 records (rows), 24 variables (columns)

Attribute Information

1. **make**: alfa-romero, audi, bmw, chevrolet, dodge, honda, isuzu, jaguar, mazda, mercedes-benz, mercury, mitsubishi, nissan, peugot, plymouth, porsche, renault, saab, subaru, toyota, volkswagen, volvo
2. **fuel-type:** diesel, gas.
3. **aspiration:** std, turbo.
4. **num-of-doors:** four, two.
5. **body-style:** hardtop, wagon, sedan, hatchback, convertible.
6. **drive-wheels:** 4wd, fwd, rwd.
7. **engine-location:** front, rear.
8. **wheel-base:** continuous from 86.6 120.9.
9. **length:** continuous from 141.1 to 208.1.
10. **width:** continuous from 60.3 to 72.3.
11. **height:** continuous from 47.8 to 59.8.
12. **curb-weight:** continuous from 1488 to 4066.
13. **engine-type:** dohc, dohcv, l, ohc, ohcf, ohcv, rotor.
14. **num-of-cylinders:** eight, five, four, six, three, twelve, two.
15. **engine-size:** continuous from 61 to 326.
16. **fuel-system:** 1bbl, 2bbl, 4bbl, idi, mfi, mpfi, spdi, spfi.
17. **bore:** continuous from 2.54 to 3.94.
18. **stroke:** continuous from 2.07 to 4.17.
19. **compression-ratio**: continuous from 7 to 23.
20. **horsepower:** continuous from 48 to 288.
21. **peak-rpm:** continuous from 4150 to 6600.
22. **city-mpg:** continuous from 13 to 49.
23. **highway-mpg**: continuous from 16 to 54.
24. **price:** continuous from 5118 to 45400.

Univariate Analysis

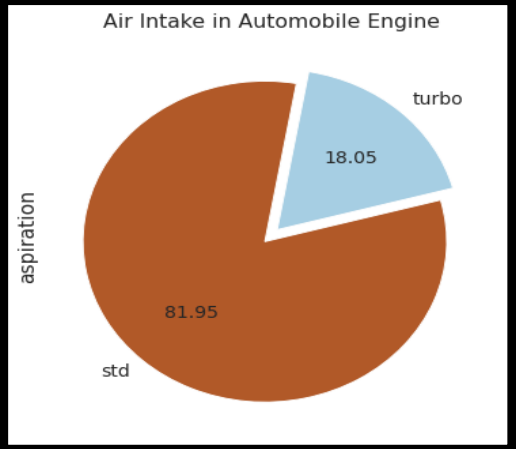


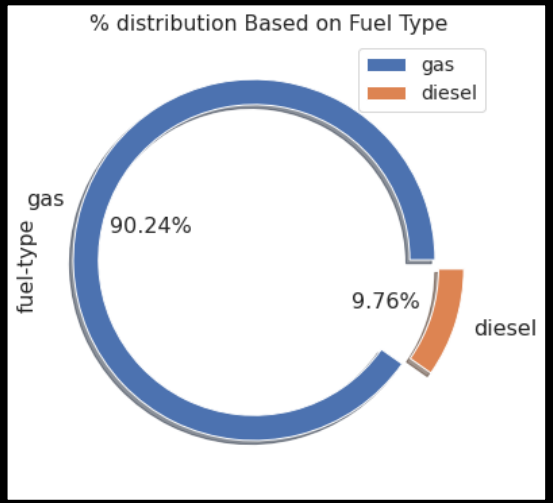
Analyzing ‘Make’ of Car we observe that

- Toyota is the most demanding make
- Toyota is followed by Nissan

From ‘Aspiration’ we observe:

- ~82 % of cars uses std. air intake
- ~18 % of cars having turbo air intake

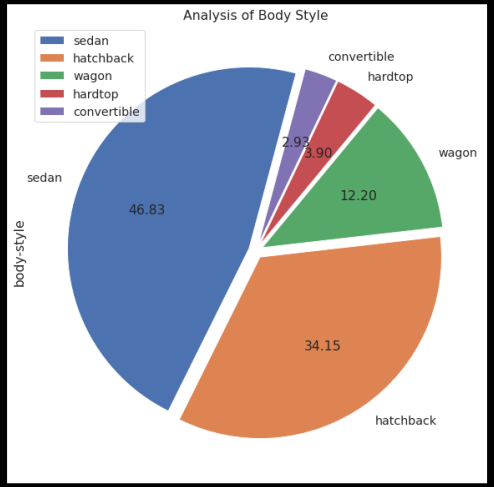




Analyzing ‘fuel-type’ we observe that :

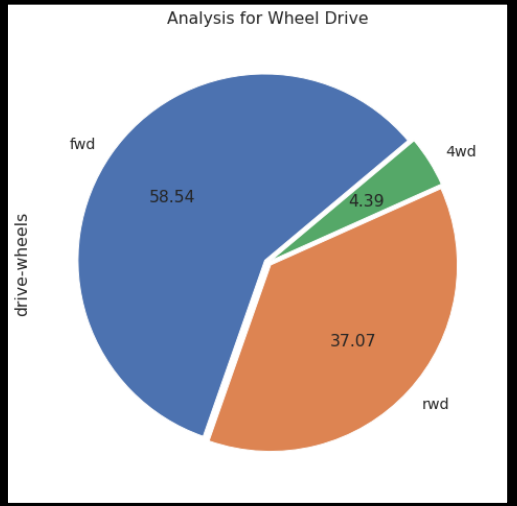
- Gas is the most preferred fuel (90.24 %)
- Very less cars uses Diesel as fuel (9.76%)

Wheel – Drive and Body – Style



If we analyze Body Style :

- Sedan is the most liked Body Style with approx. value of 47 %
- Convertibles are the least demanding Body style with approx. 3 %



With Wheel Drive we can say that:

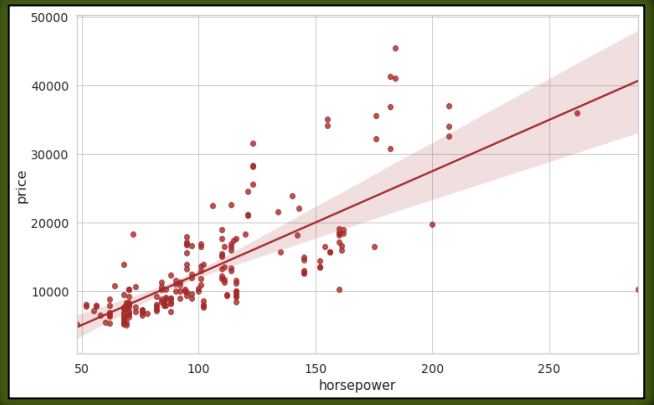
- Most of the cars uses Front Wheel Drive (fwd)
- Only 4 % of cars uses 4 wheel drive (4wd), so is least demanding

Summary of Univariate Analysis

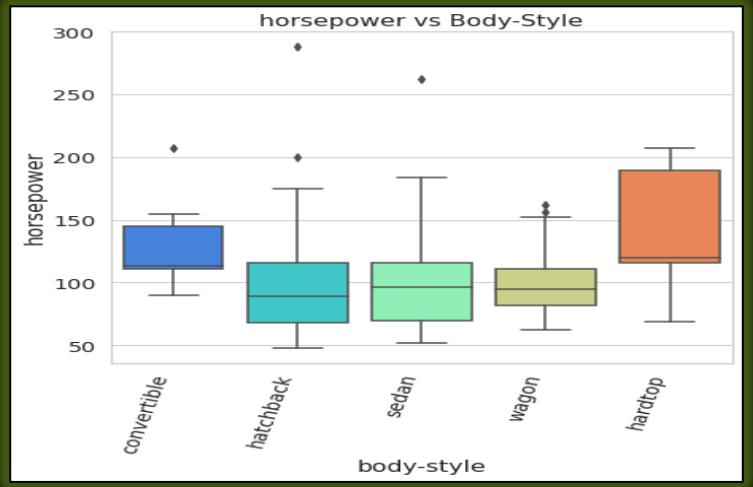
From Univariate Analysis we observe that:

- In Make , Toyota is the most demanding car (~21 %) and Dodge is least demanding (~6 %) ,
- Among STD and Turbo aspiration, Standard (std.)is used mostly ,
- 90 % of cars uses Gas as fuel over Diesel (~10%),
- Sedan body-style is highest among all other styles available
- Maximum nos. of cars uses Front Wheel Drive (fwd)

Horsepower Vs Price Analysis



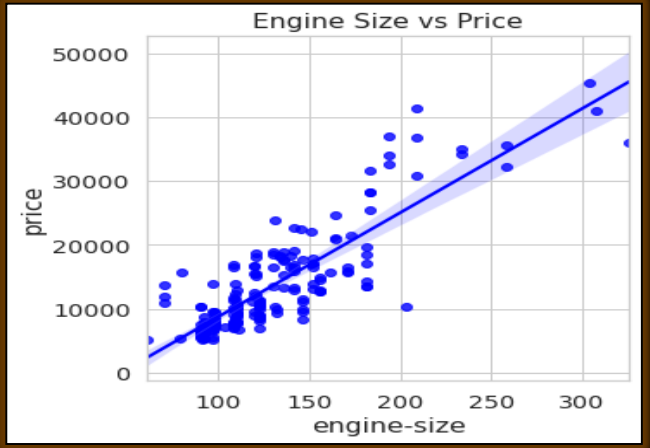
Car pricing maintains strong positive correlation with the engine Horsepower.



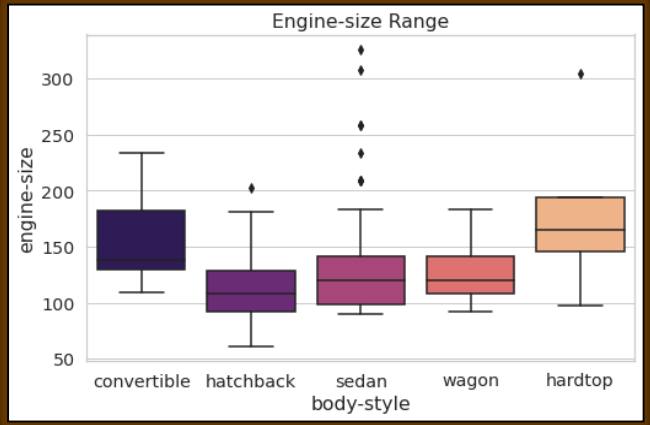
With the box plot we can say that:

- Almost in all body styles low HP engines are available
- In Hardtop, most of the vehicles found in higher HP
- In all styles, except hardtop few cars comes with high HP engines

Pricing VS Engine Size Analysis



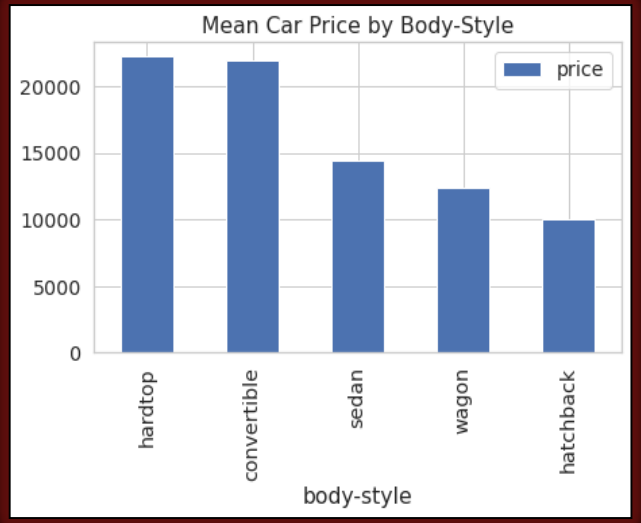
Car pricing maintains strong positive correlation with its engine size.



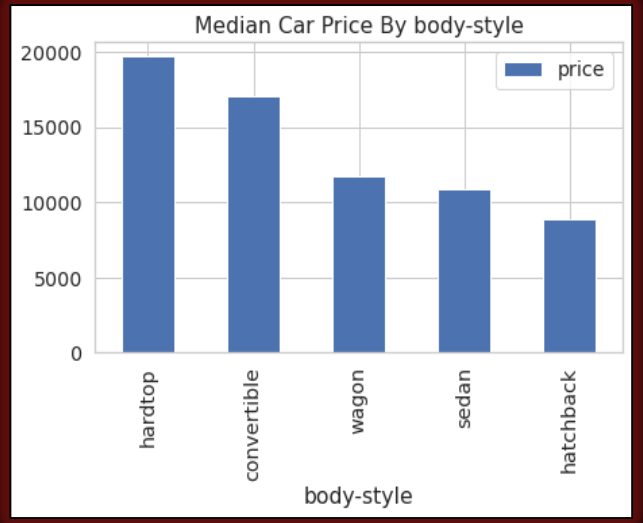
Analysis for Engine-size vs Body-style :

- Convertibles have engine size available in higher ranges,
- In Sedan Class we observe some vehicles with excess engine size (seems possible outliers)

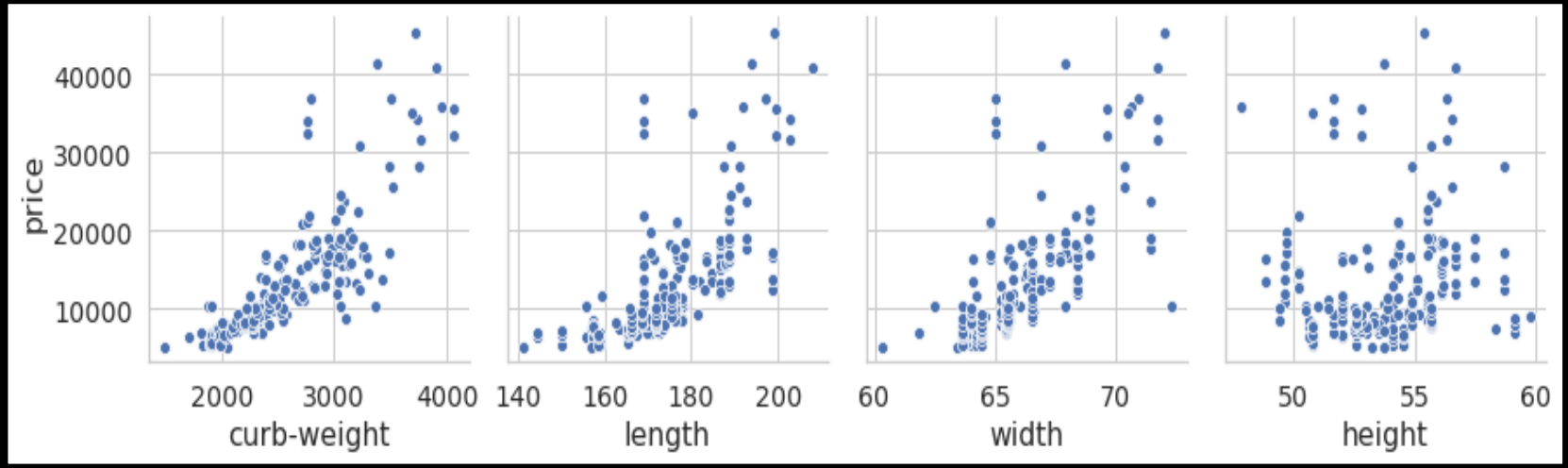
Body Styles and Pricing



Convertibles and hardtops are the Costliest car models



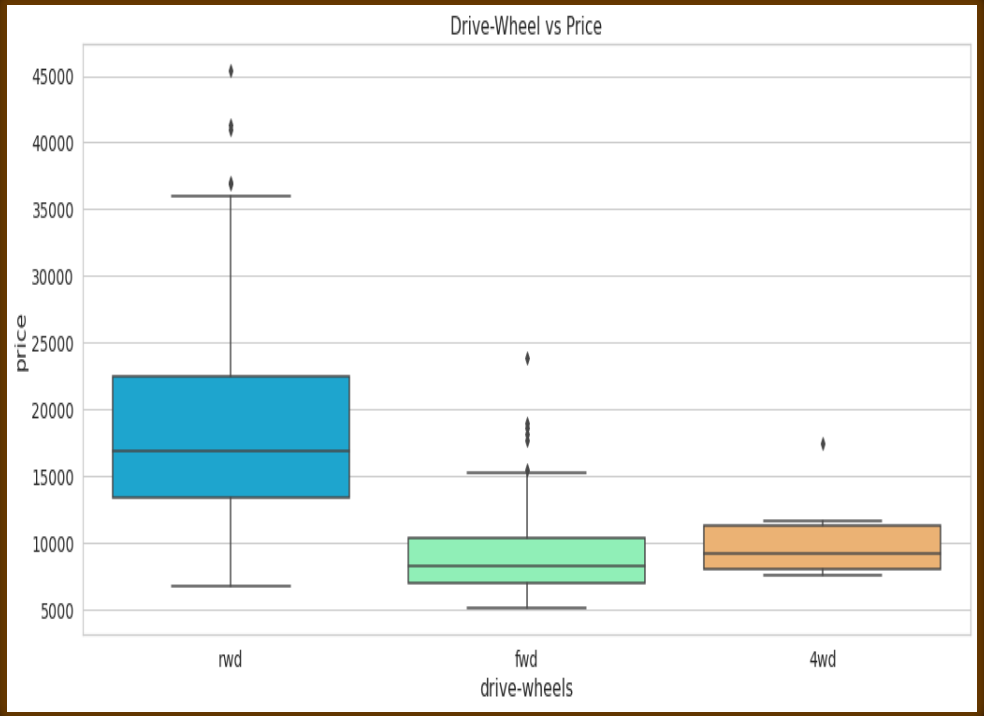
Body Size ,curb-weight vs Pricing



Based on above plot we analyses that :

- Car's length and width have strong positive correlation with its pricing
- However height doesn't have any impact on Car's pricing
- Curb-weight is also positively correlated with pricing of car's

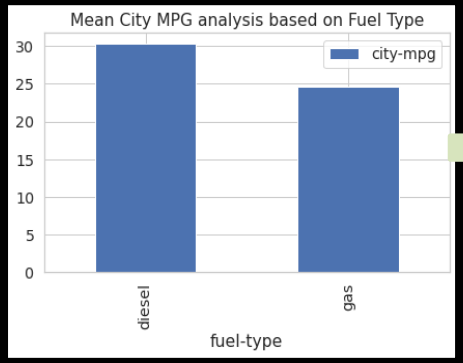
Wheel Drive vs Pricing



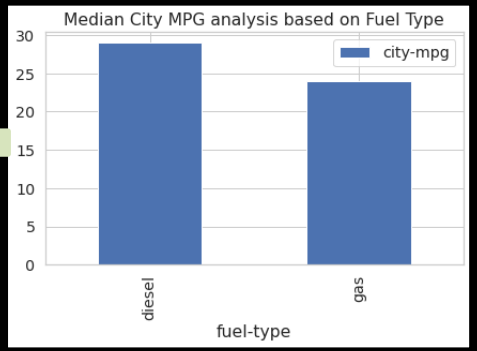
This analysis shows that:

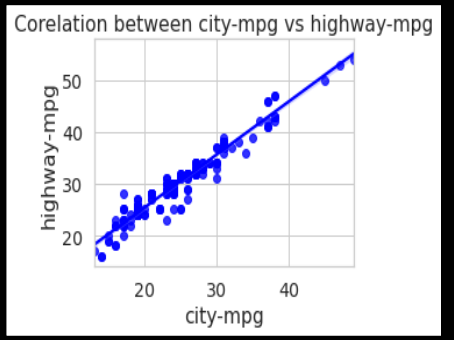
- rwd car's available in almost all price range
- fwd and 4wd are available only in low price range with some outliers available,
- 4wd car's available in lowest price slot among all

Fuel-efficiency vs Fuel type vs Pricing



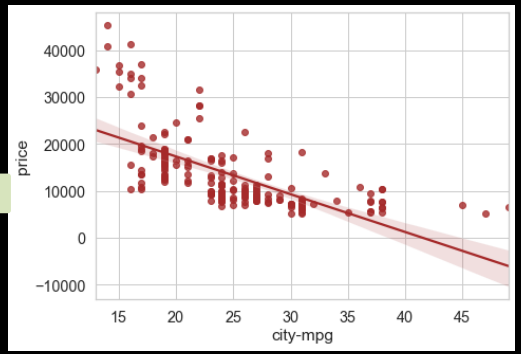
Diesel cars have **good city-mpg** as compare to **Gas**





city-mpg is strongly correlated with highway-mpg, so we are using city-mpg for our analysis

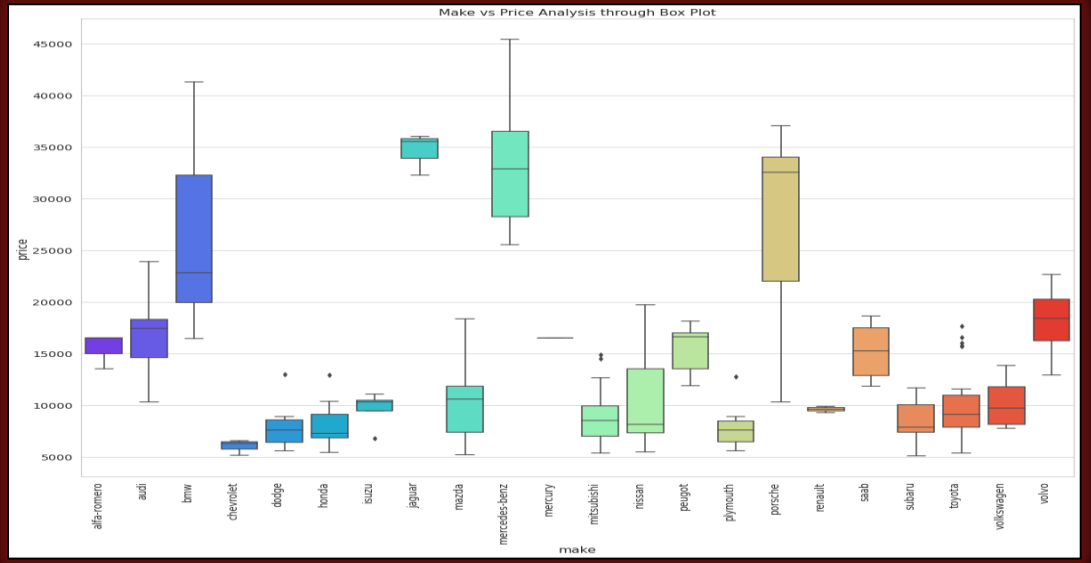
Fuel efficiency shows negative correlation with pricing,



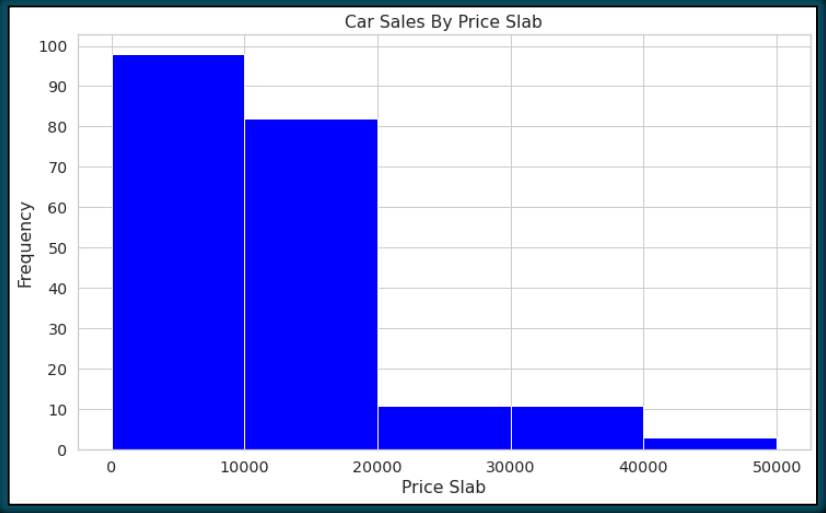
Make vs Pricing

This analysis shows that:

- Price of BMW, Mercedes-Benz, Jaguar and Porsche are very high,
- These models comes under premium categories.



Car Price – Data Distribution



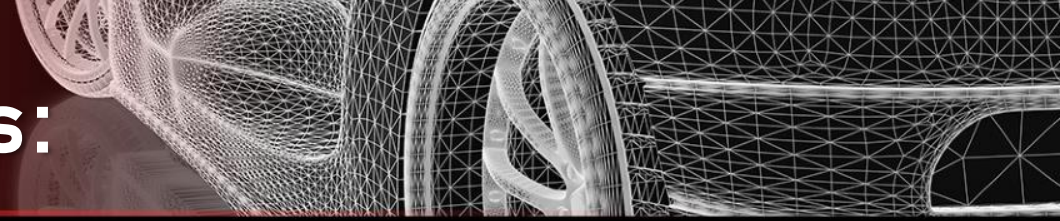
Distribution of Price

Majority of cars belong to the lower price brackets (< 20K) even though there are cars that go up to 45K

Conclusion Based on Analysis

- Diesel cars produce better mileage compare to cars with Gas fuel type.,
- Most of the car's uses Low Horse power engine to keep price low.
- Curb-weight have positive effect of length and width
- Price of car's goes up if size and curb-weight increases
- Price vs Make analysis shows that BMW, Jaguar , Porsche and Mercedes-Benz belongs to premium range due to high cost
- Majority of cars belong to the lower price brackets; i.e. < 20K
- Price get impacted with increase in size and curb-weight as they are positively correlated

Actionable Insights:



- ✓ Cars having smaller engines and less horse power produce better mileage.
- ✓ To sell more cars we have to price it under 20,000/-,
- ✓ Hatchback with Diesel as fuel will be preferred over any other combination as it gives good city average,
- ✓ To make price low we can use Low power engine,
- ✓ Engines with Low width and length will be used to minimize the price

THANK YOU