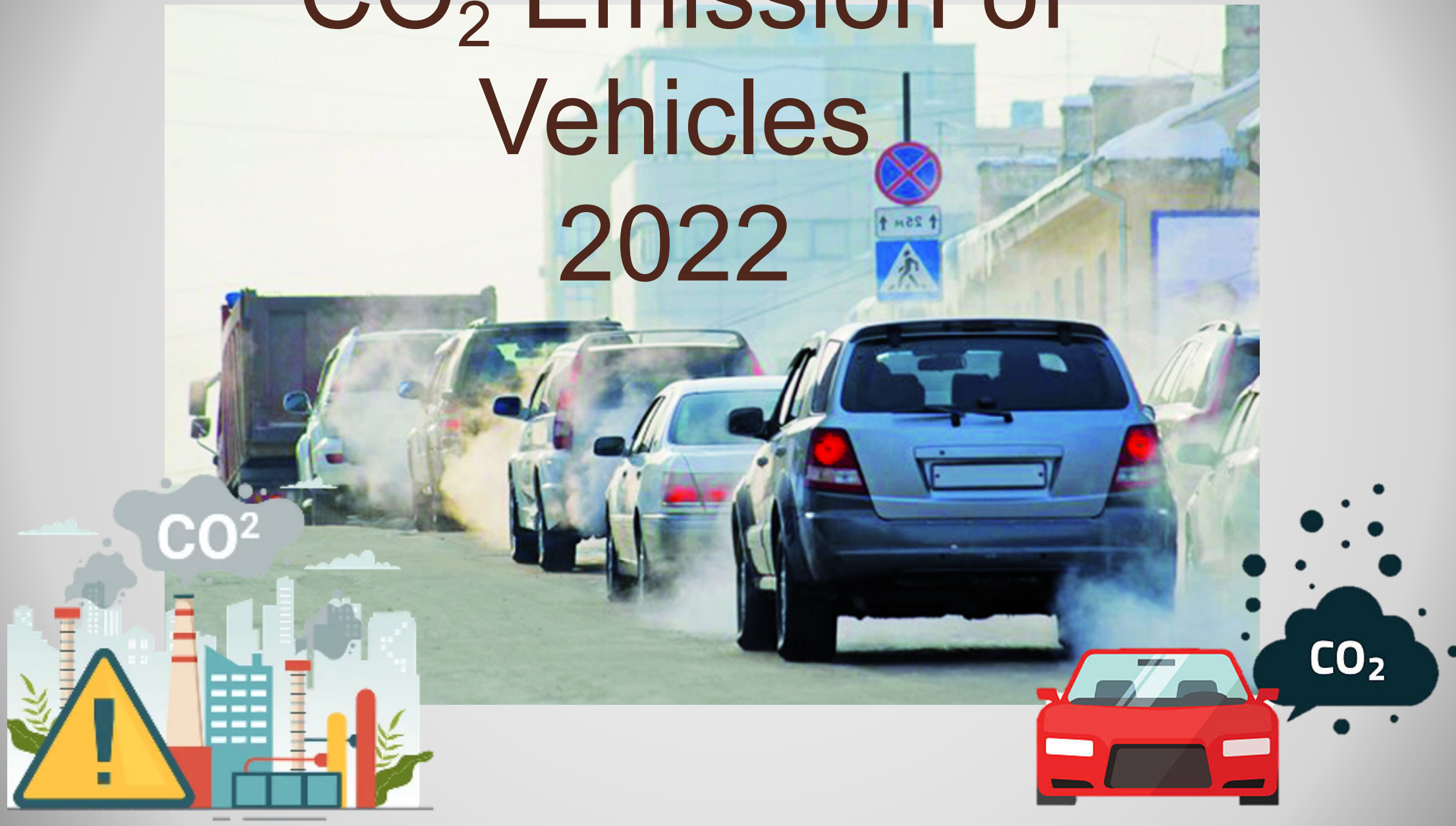
CO_2 Emission of Vehicles 2022



Group Members

- M. P. I Malinda
- T. D. T Weerasighe
- E. M. K. C Ekanayake
- L.A. P Kavinda
- J. H. R. M Jayasundara
- M. H. H. D. N Prematilaka
- G. R. I. S Gunasekara
- S. A. V Samaraweera
- M. A. P Peshan

Outline

1. Introduction
2. Objectives
3. Methodology
4. Results And Discussion
5. Conclusion
6. References

Introduction

- We know that human is the most intelligent creature in the whole world.
- So, day by day, generation to generation human was developed and, they invented lot of things with the help of technology.
- As a result of that there happened a vast progress in technology and industrialize.
- With that progress parallelly occurred the environmental pollution too since human missed their close friend, environment. Because of that mainly in this era we’re being facing many challenges. The foremost challenge among

Objectives

- To analyze the factors associated with green cars which have low CO_2 emissions.
- To identify the better vehicles to use in city roads, highways, and combination.
- To analysis the which model of vehicles have the least and most fuel consumption.
- To analysis the which type of fuel emits the least amount of CO_2
- To suggest the solutions for reduce CO_2 emission of vehicles.

Methodology

- Graphical Analysis Part
- Regression Analysis Part
  - Data Source: Kaggle

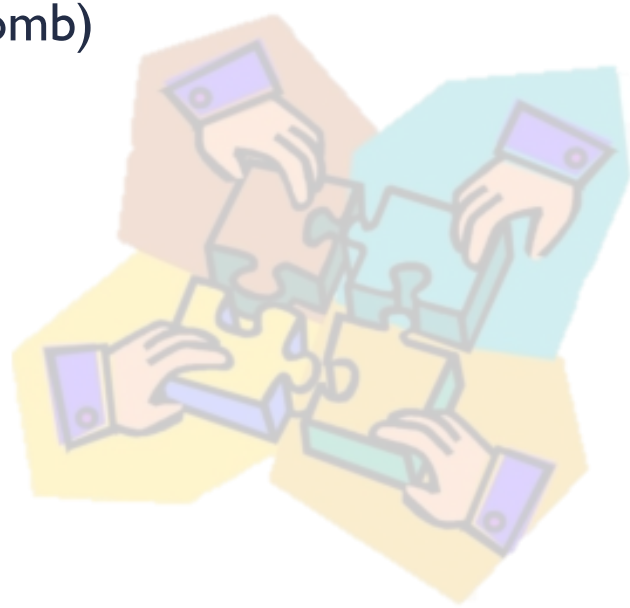
There are one dependent and five number of continuous predictor variables get for the regression part

Graphical Analysis Part

- Number of Cylinders
- Engine-Size
- Fuel-Consumption (City)
- Fuel-Consumption (Hwy)
- Fuel-Consumption (Comb)
- Make
- Model
- Vehicle-Class
- Fuel-Type
- Transmission

Regression analysis Part

- Y_1 - CO_2 emission
- X_1 - Engine size
- X_2 - Number of cylinders
- X_3 - Fuel consumption (City)
- X_4 - Fuel consumption (HWY)
- X_5 - Fuel consumption (comb)



Description of variables

- Model-Year: Year of Published
- Make: Brands of Vehicles
- Model: Vehicles Models
- Vehicle-Class: Sizes of the vehicles
- Engine-Size: vehicle engine size

Nature of Variables :

- Categorical variables:- Model-Year, Make, Model, Vehicle-Class, Fuel-Type, Transmission

- Quantitative variables: - Cylinders, Engine-Size, Fuel-Consumption(City), Fuel-Consumption(Hwy), Fuel-Consumption(Comb), CO2-Emission

Graphical Analysis Part

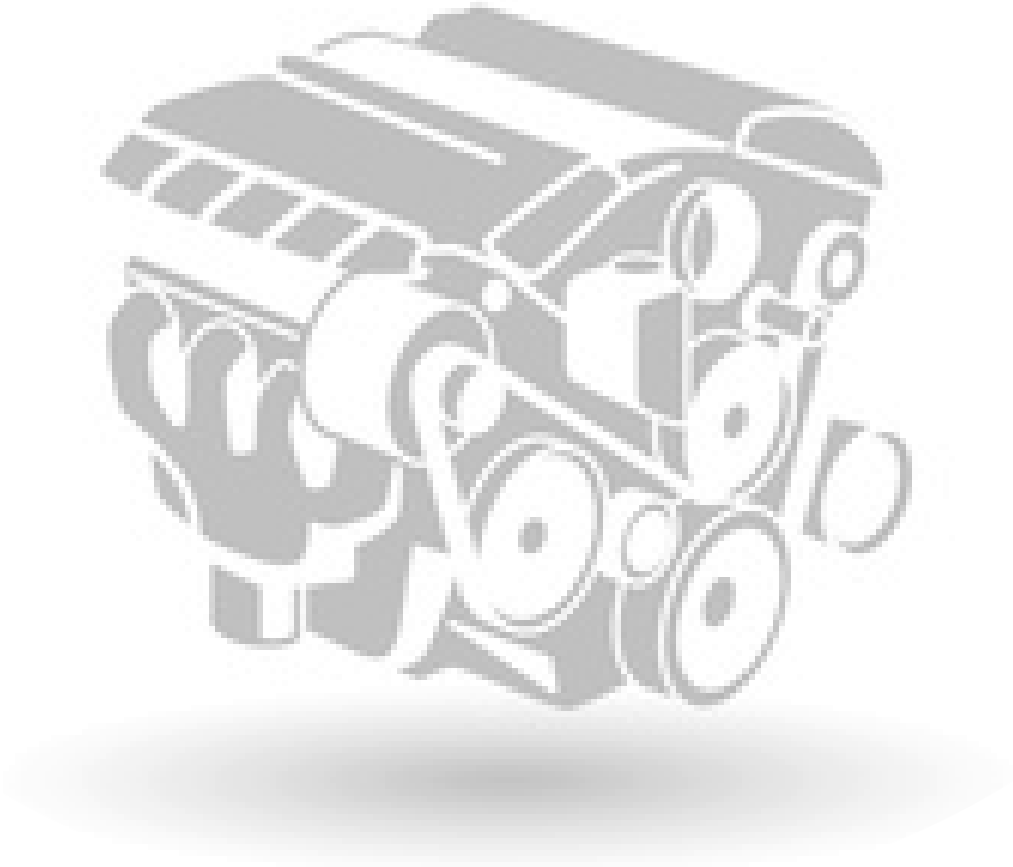
In graphical analysis part

- A fitted line plot was used to identify the variation of engine size with CO_2 emission.
- An individual value plot was used to identify the variation of CO_2 emission with number of cylinders.
- another bow plot was used to identify the variation of CO_2 emission with fuel type.
- A pie chart was used to identify the usage of fuel types.
- In other analysis part, box plots were used to identify the better vehicles to use in city, highway and combination.

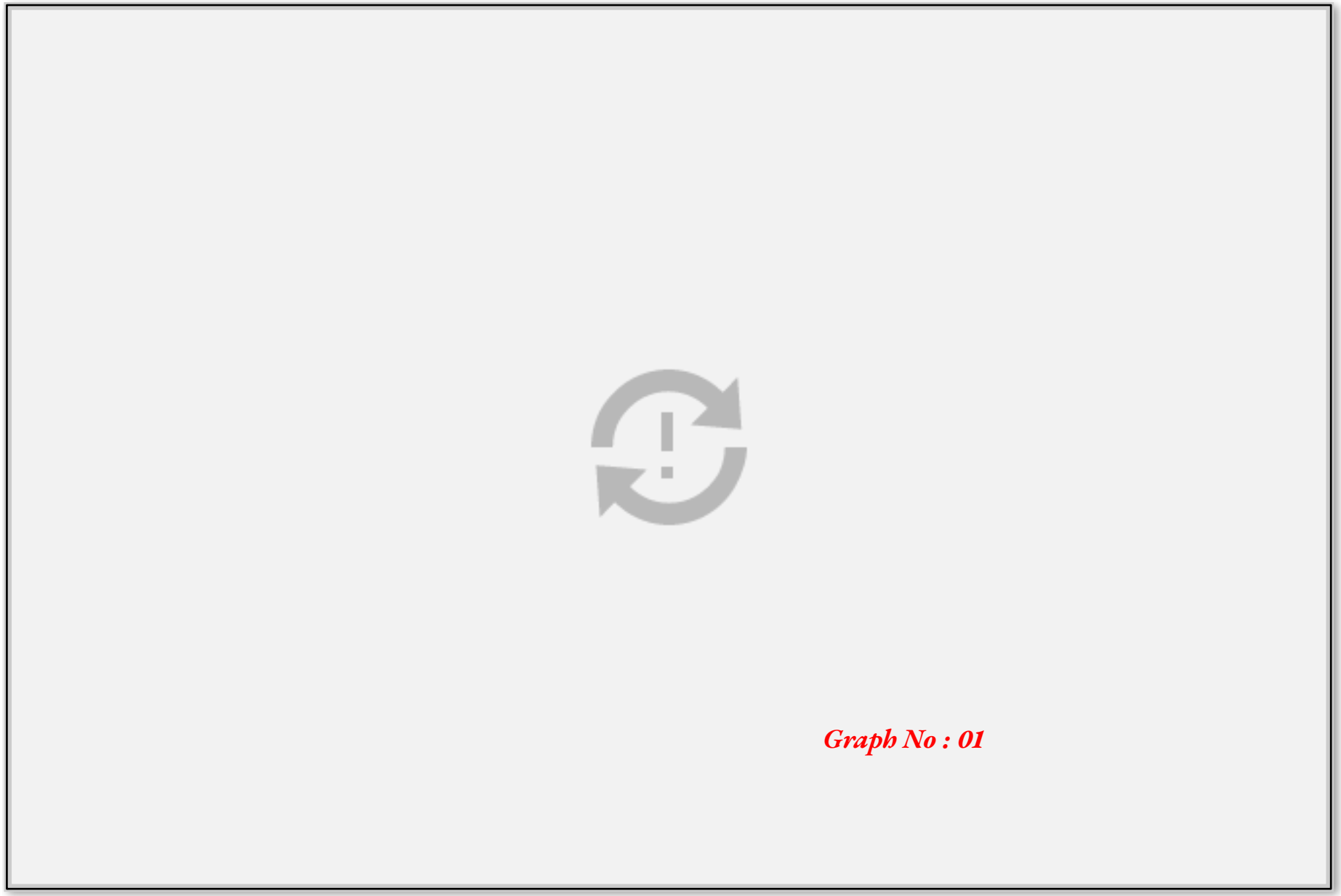
❖ Factors associated with green cars which have low CO2 emissions…

we found out 4 main factors that associated with green cars which have low CO2 emissions.

1. Engine size
2. Number of Cylinders
3. Fuel type
4. Fuel consumption

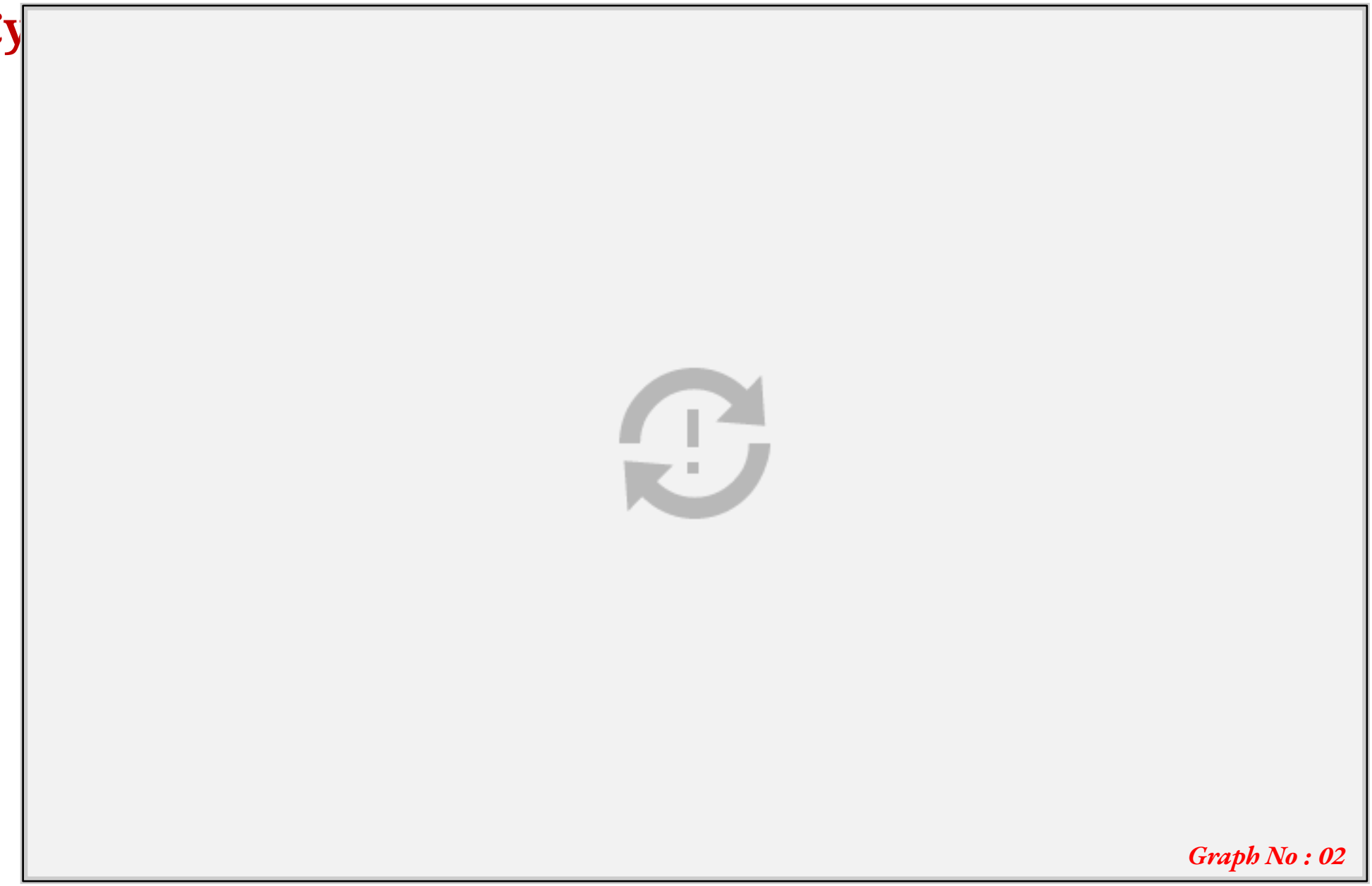


1. Engine Size



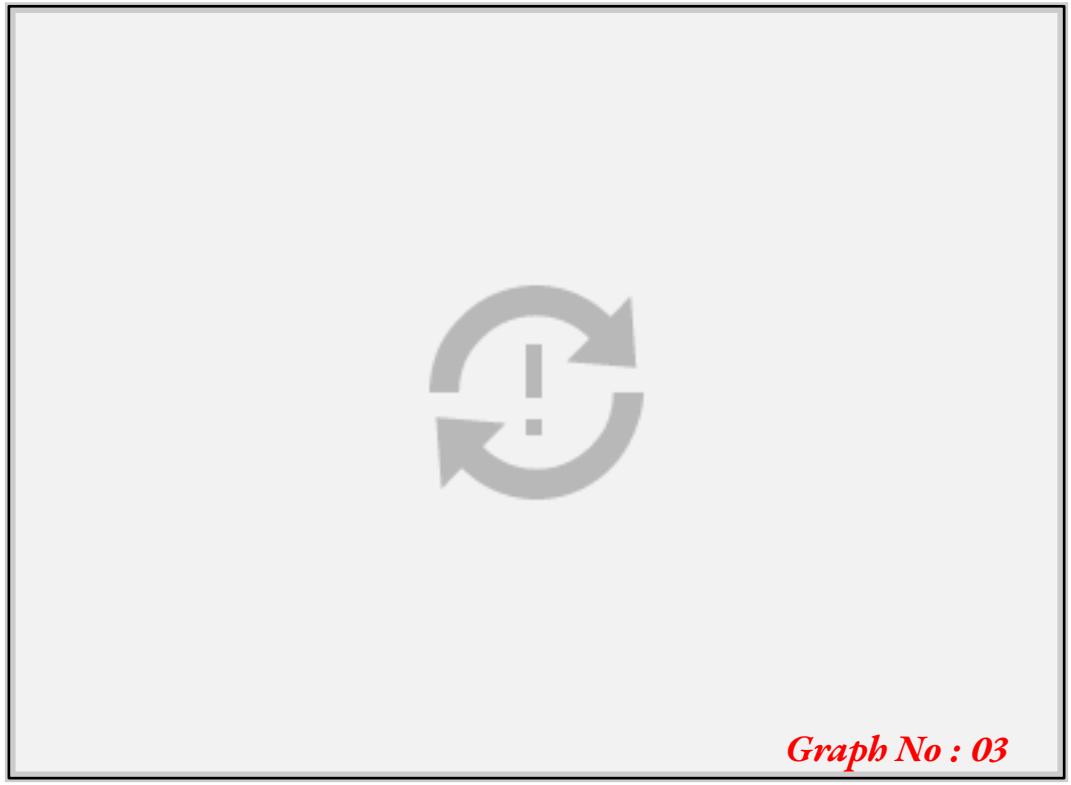
Graph No : 01

2. Number of Cy

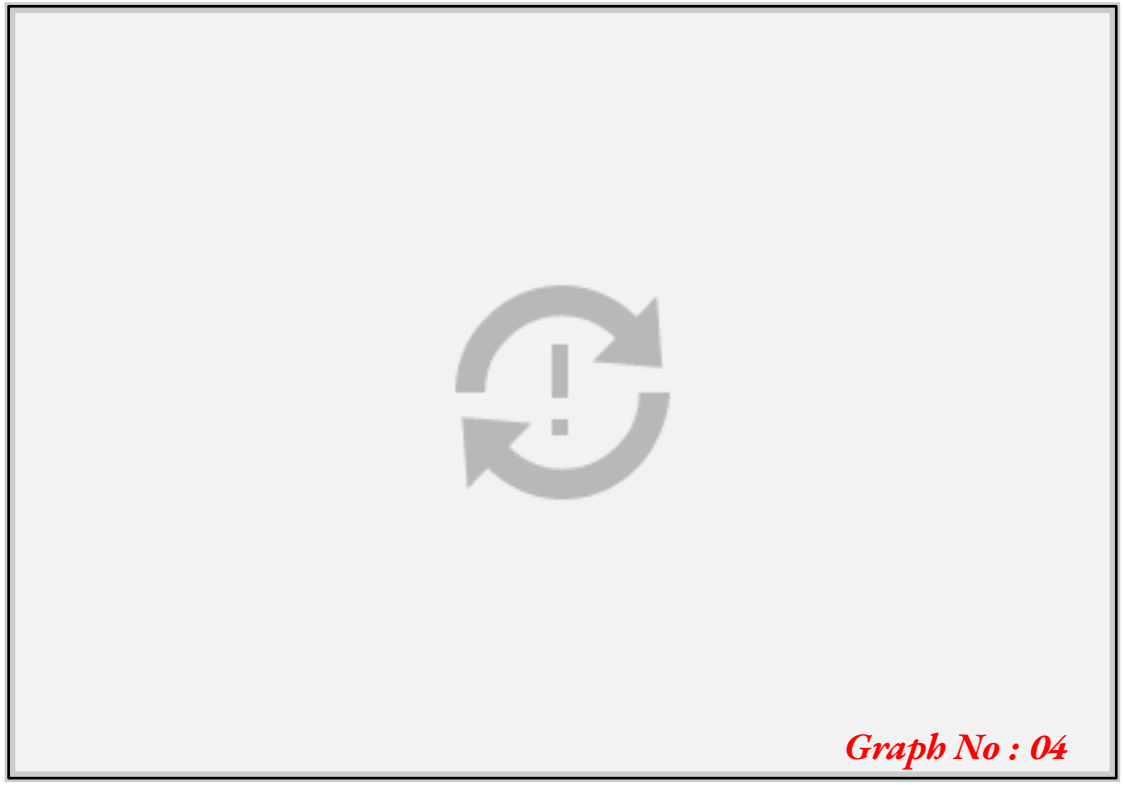


Graph No : 02

3. Fuel Type



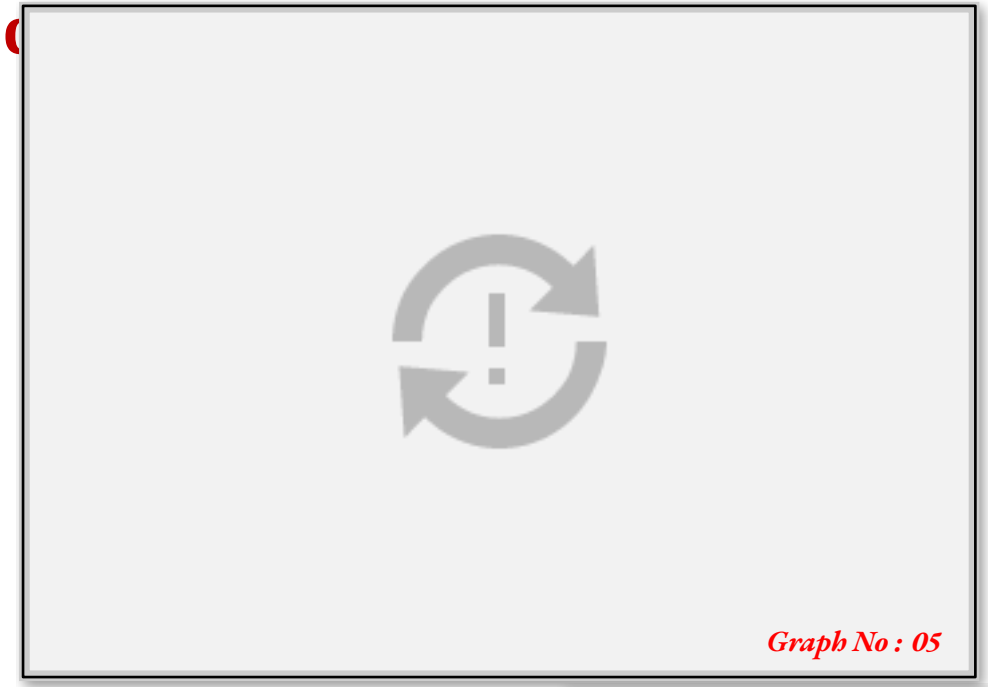
Graph No : 03



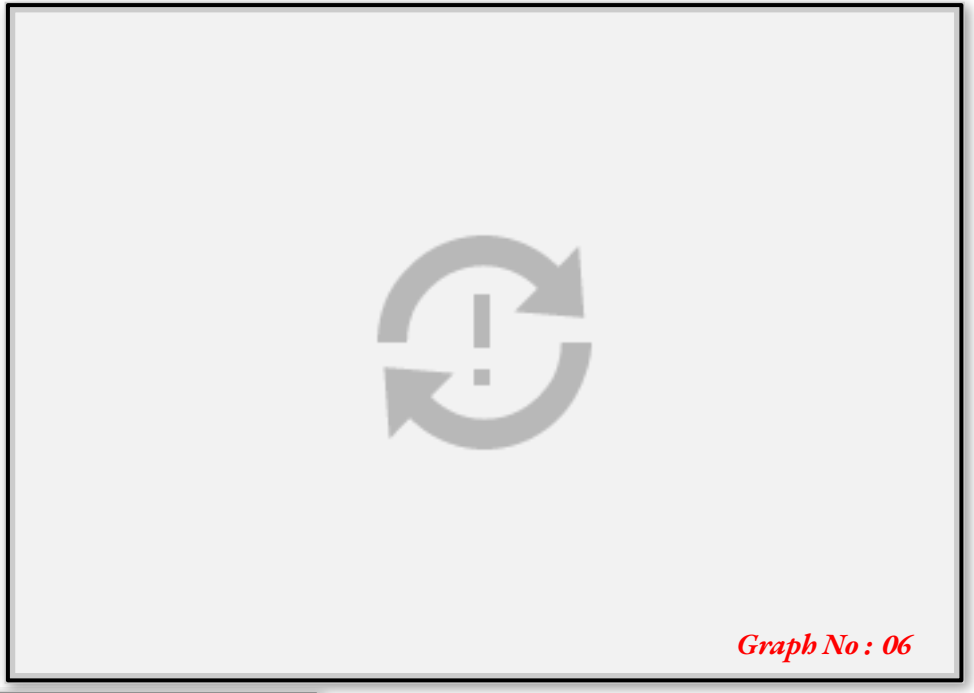
Graph No : 04

Variable	Fuel Type	Count	Mean	Median
CO_2 emission	D	28	271.57	272.00
	E	14	294.86	299.50
	X	446	237.58	232.00
	Z	458	278.35	273.50

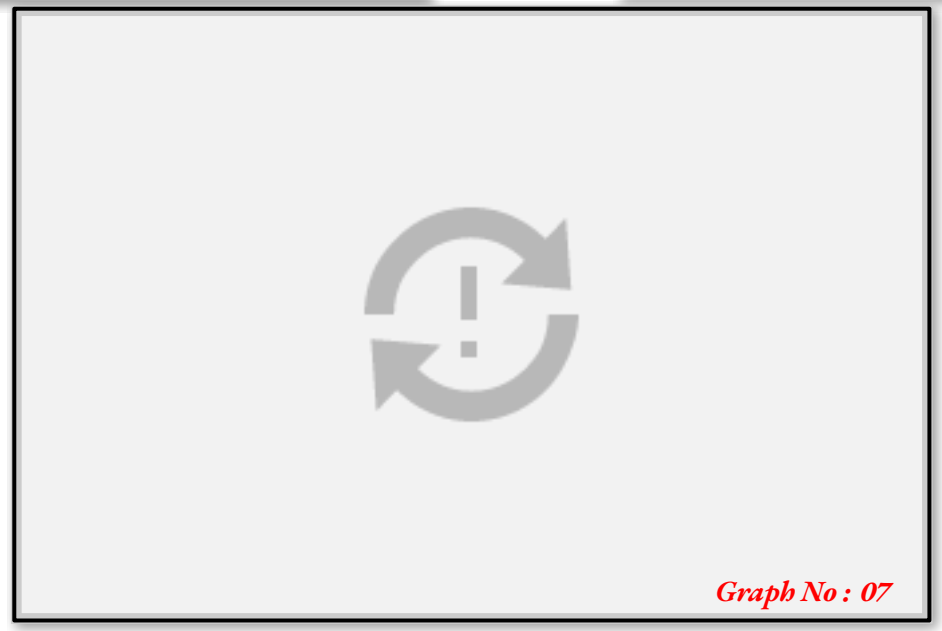
4. Fuel



Graph No : 05



Graph No : 06

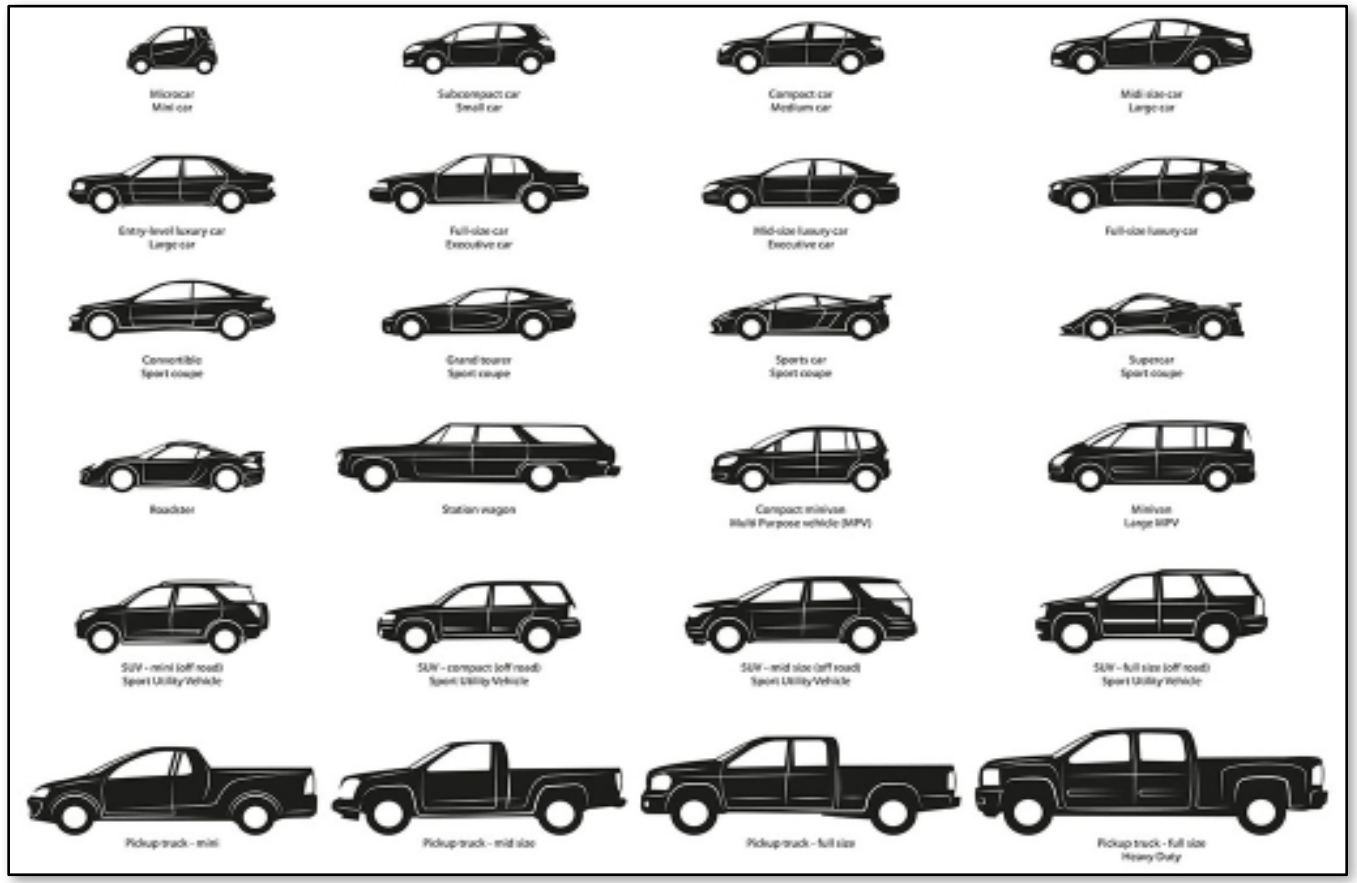


Graph No : 07

Other Analysis Part

- **Identifying the better vehicles to use in city, highway and combination**

To analysis this scenario, firstly we divided our data set into vehicles class wise. People buying vehicles for their own choices and needed.



Measurement of fuel consumption is **Litter per km (L/100km)**

1. Mini compact

Graph No : 08

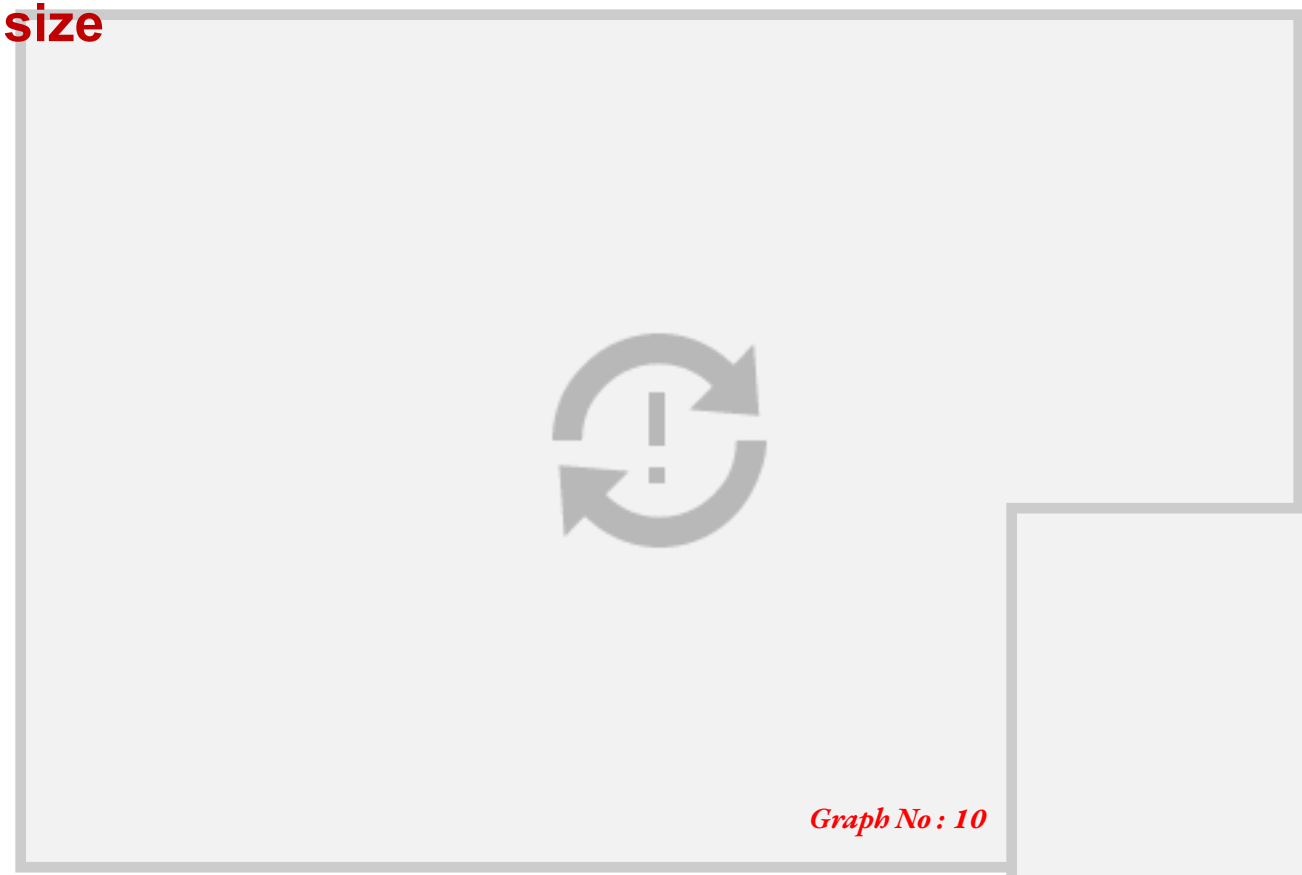
MINI, Subaru and Toyota vehicle's fuel consumption is low on city, highway and combination while **Bentley** has the highest fuel consumption. But these vehicles will be used less than 18 litter per 100km.

2. Special Purpose Vehicles

Graph No : 09

Less Fuel consumption in the brand Ford in City roads and Combination roads. In highway roads Ram vehicles are better than Ford vehicles. Mercedes - Benze is the highest fuel consumption Vehicle in special purpose vehicle.

3. Station Wagon : Small size

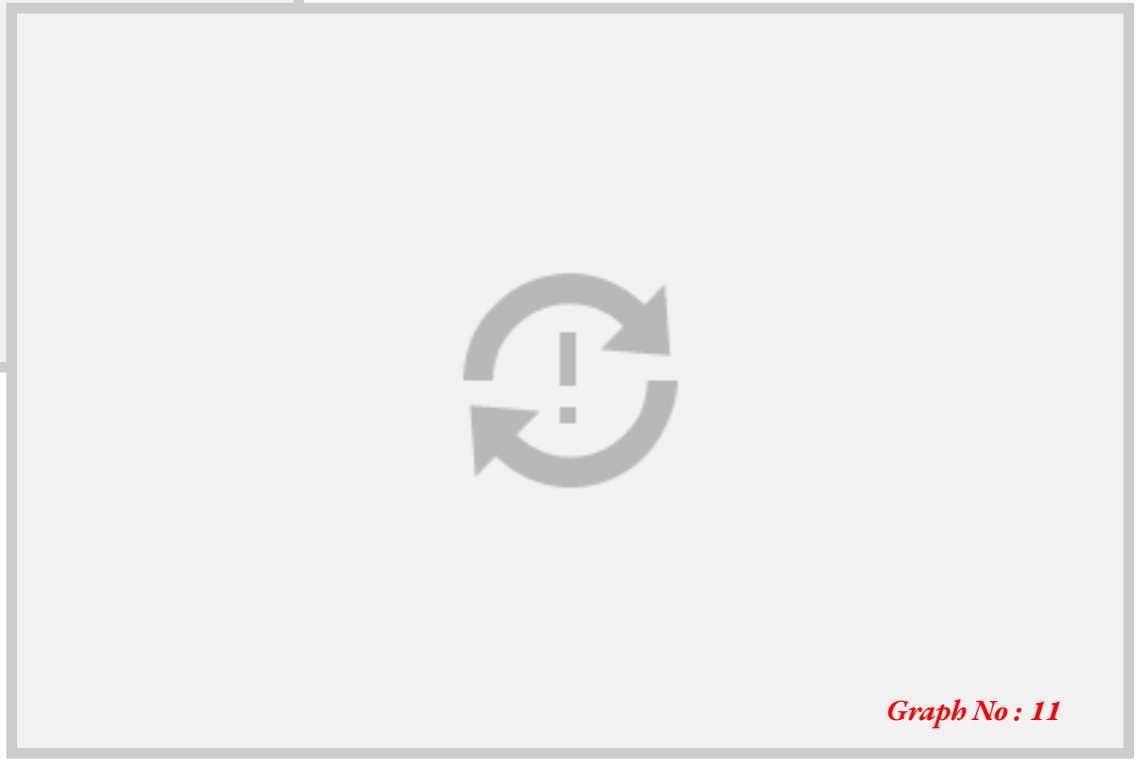


Graph No : 10

Mercedes - Benze has the highest Fuel consumption.

Least fuel consumption on City, Highway, and Combination roads vehicle models of KIA

- **Niro**
- **Niro FE**



Graph No : 11

4. Pickup Truck : Standard

Graph No : 12

Chevrolet is the highest fuel consumption Vehicle in pickup trucks (standard).

- **Silverado 4WD Mud Terrain Tire (No Stop-Start) – In City/ Combination**
- **Silverado 4WD Mud Terrain Tire FFV – In Highway**

Comparing to the other, less Fuel consumption in the brand Honda in City roads, Highway and Combination also.

Graph No : 13

Regression Analysis Part

In regression analysis part

- First, we use stepwise selection method to select best regression model.
- We check model assumption. If our assumption is not satisfying, then we hope to transform our variables.
- R^2 value is used to check the accuracy of our model.
- Pearson correlation values are used to test the functional relationship between independent variable and dependent variables.
- P-value (by using ANOVA table) is used to check our predictor(x) variables significantly associated with the response(y) variable at 5% significant level.

- **Using step-wise selection method select the best model consider the all 5 predictor variables.**

Regression Analysis: CO2 Emission versus Engine Size(, Cylinders, Fuel Consump, ...

Stepwise Selection of Terms

Candidate terms: Engine Size(L), Cylinders, Fuel Consumption (City (L/100 k, Fuel Consumption(Hwy (L/100 km), Fuel Consumption(Comb (L/100 km

	-----Step 1----		-----Step 2----		-----Step 3----	
	Coef	P	Coef	P	Coef	P
Constant	33.72		19.68		19.09	
Fuel Consumption (City (L/100 k	18.027	0.000	13.049	0.000	10.842	0.000
Fuel Consumption(Hwy (L/100 km)			8.147	0.000	9.012	0.000
Cylinders					3.546	0.000
S	16.7584		15.1530		14.7315	
R-sq	93.24%		94.48%		94.79%	
R-sq(adj)	93.24%		94.47%		94.77%	
R-sq(pred)	93.18%		94.37%		94.66%	
Mallows' Cp	278.12		56.50		2.83	

α to enter = 0.15, α to remove = 0.15

According to the step-wise selection method, If a model given,

- **The highest R^2 - value**
- **Highest adjusted R^2 - value**
- **The lowest Mallon's C, statistics value**
- **All P - value < 0.05,**

then we can say that the model is the best model. So, according to the above output step no 3 gives the best model. Because all above conditions are satisfied.

Number of Cylinders

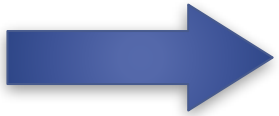
Engine Size

Fuel-Consumption (City)

Fuel-Consumption (Hwy)

Fuel-Consumption (Comb)

Numerical Variables



Number of Cylinders

Fuel-Consumption (City)

Fuel-Consumption (Hwy)

Best Selected Model

Check the basic assumptions for best selected model

Assumption	Graphical representation	Method	Satisfied or Not satisfied
Normality	**Normal probability plot**	The points on the probability plot should lie in an approximately straight diagonal line	**Not perfectly satisfied**
Linearity	**Scalar matrix plot**	The point on the plot should be fall along a straight line	**Not perfectly satisfied**
Independency	**Durbin – Watson value**	The Durbin – Watson value should be between 1 and 3	**Durbin – Watson value is 1.21769 (Satisfied)**
Homoscedasticity	**Residual & Fitted plot**	The residuals appear to be randomly around zero	**Not satisfied**

After checking the assumptions, there were some outliers in our graphs. So, we removed those points using MINITAB software.

- **After removing outliners, step-wise selection method select the best model consider the all 5 predictor variables.**

Regression Analysis: CO2 Emission versus Engine Size(, Cylinders, Fuel Consump, ...

Stepwise Selection of Terms

Candidate terms: Engine Size(L), Cylinders, Fuel Consumption (City (L/100 k, Fuel Consumption(Hwy (L/100 km), Fuel Consumption(Comb (L/100 km

	-----Step 1-----		-----Step 2-----		-----Step 3----	
	Coef	P	Coef	P	Coef	P
Constant	18.64		-0.031		-0.019	
Fuel Consumption (City (L/100 k	19.275	0.000	12.7439	0.000	8.038	0.000
Fuel Consumption(Hwy (L/100 km)			10.7321	0.000	6.869	0.000
Fuel Consumption(Comb (L/100 km					8.57	0.000
Engine Size(L)						
Cylinders						
S	9.28556		1.37919		1.34157	
R-sq	97.70%		99.95%		99.95%	
R-sq(adj)	97.69%		99.95%		99.95%	
R-sq(pred)	97.68%		99.95%		99.95%	
Mallows' Cp	42064.25		73.22		23.28	

	-----Step 4----		-----Step 5-----	
	Coef	P	Coef	P
Constant	0.155		0.425	
Fuel Consumption (City (L/100 k	7.861	0.000	7.957	0.000
Fuel Consumption(Hwy (L/100 km)	6.789	0.000	6.779	0.000
Fuel Consumption(Comb (L/100 km	8.77	0.000	8.71	0.000
Engine Size(L)	0.1545	0.011	0.4088	0.000
Cylinders			-0.2626	0.000
S	1.33742		1.32703	
R-sq	99.95%		99.95%	
R-sq(adj)	99.95%		99.95%	
R-sq(pred)	99.95%		99.95%	
Mallows' Cp	18.74		6.00	

According to this output step no 5 gives the best model. Because all above conditions are satisfied.

- **The highest R^2 - value**
- **Highest adjusted R^2 – value**
- **The lowest Mallon's C, statistics value**
- **All P-value < 0.05,**

Check the basic assumptions for second best selected model

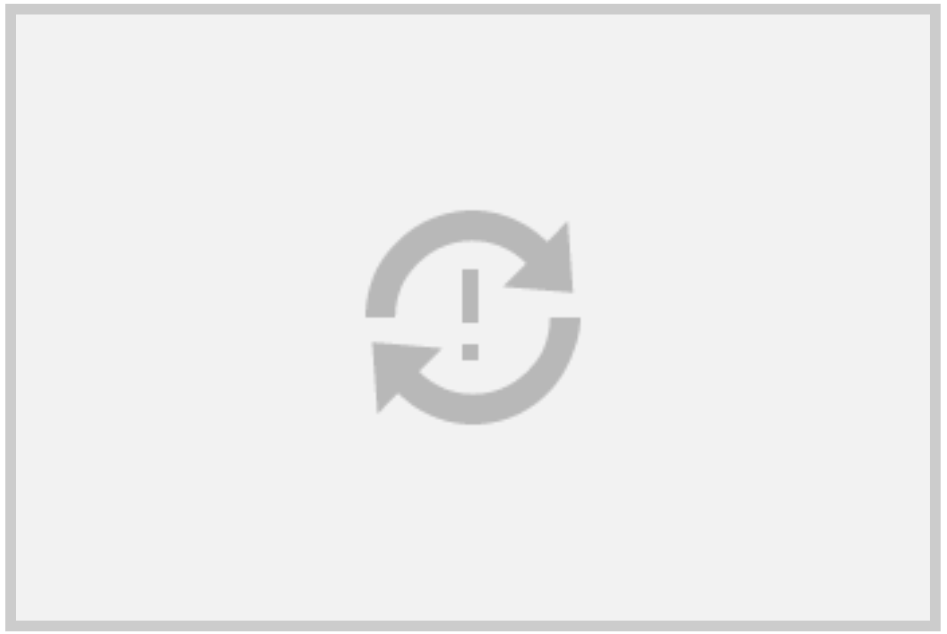
Assumption	Graphical representation	Method	Satisfied or Not satisfied
Normality	**Normal probability plot**	The points on the probability plot should lie in an approximately straight diagonal line	**Not satisfied**
Linearity	**Scalar matrix plot**	The point on the plot should be fall along a straight line	**Not perfectly satisfied**
Independency	**Durbin – Watson value**	The Durbin – Watson value should be between 1 and 3	**Durbin – Watson value is 1.09776 (Satisfied)**
Homoscedasticity	**Residual & Fitted plot**	The residuals appear to be randomly round zero	**Not satisfied**

Some assumptions are not satisfied. So, we used transformation methods to select the best model which is satisfied the all assumptions perfectly

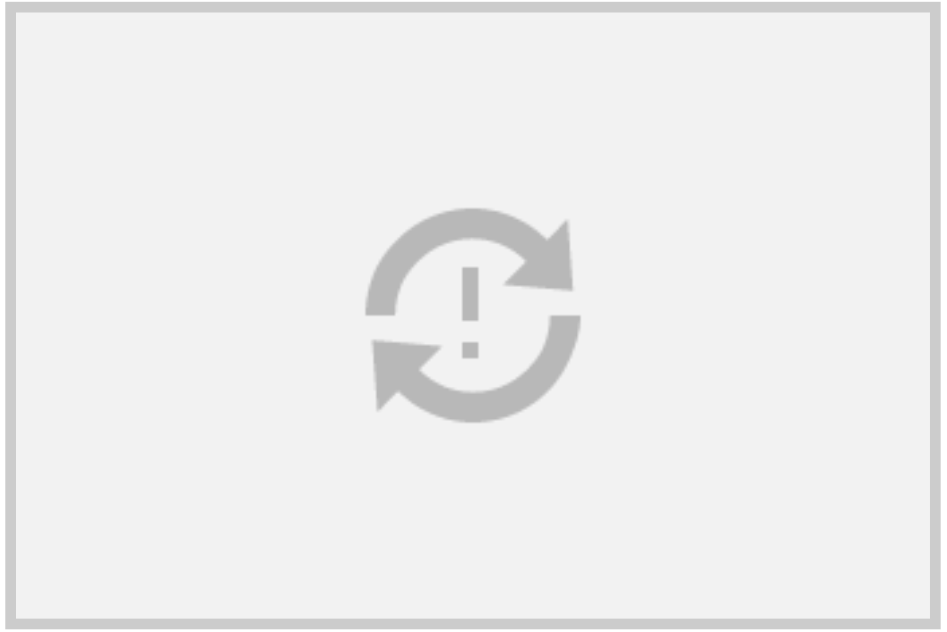
Response Variable	Predictor Variables (After transformed method)		
	Original (x)	1/x	Log(x)
Original (y)	R^2 value is 99.95% D.W. value = 1.09776 Normality and homoscedasticity not satisfy perfectly	R^2 value is 93.49% D.W. value = 1.27007 Normality not satisfy perfectly	R^2 value is 98.23% D.W. value = 1.33119 Normality not satisfy perfectly
1/y	R^2 value is 90.43% D.W. value = 1.22545 Normality, Linearity and homoscedasticity not satisfy perfectly	R^2 value is 99.95% D.W. value = 1.09734 Normality not satisfy perfectly	R^2 value is 97.73% D.W. value = 1.31207 Normality, Linearity and homoscedasticity not satisfy perfectly
Log(y)	R^2 value is 97.83% D.W. value = 1.22302 Normality not satisfy perfectly	R^2 value is 98.05% D.W. value = 1.30320 homoscedasticity not satisfy perfectly	R^2 value is 99.95% D.W. value = 1.00308 All the assumptions are satisfied perfectly

Results for check the normality satisfaction

Probability Plot for best selected predictors

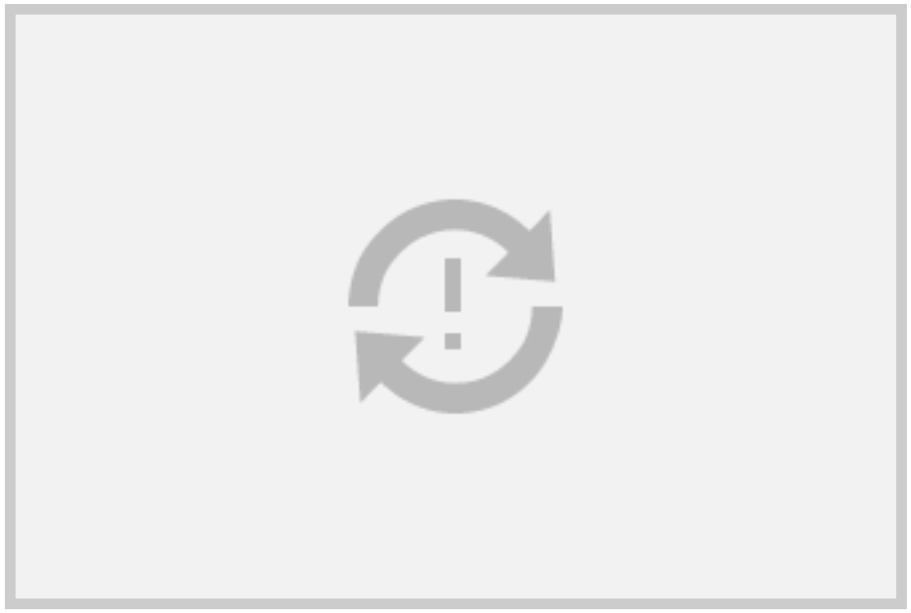


Probability Plot for Log predictors

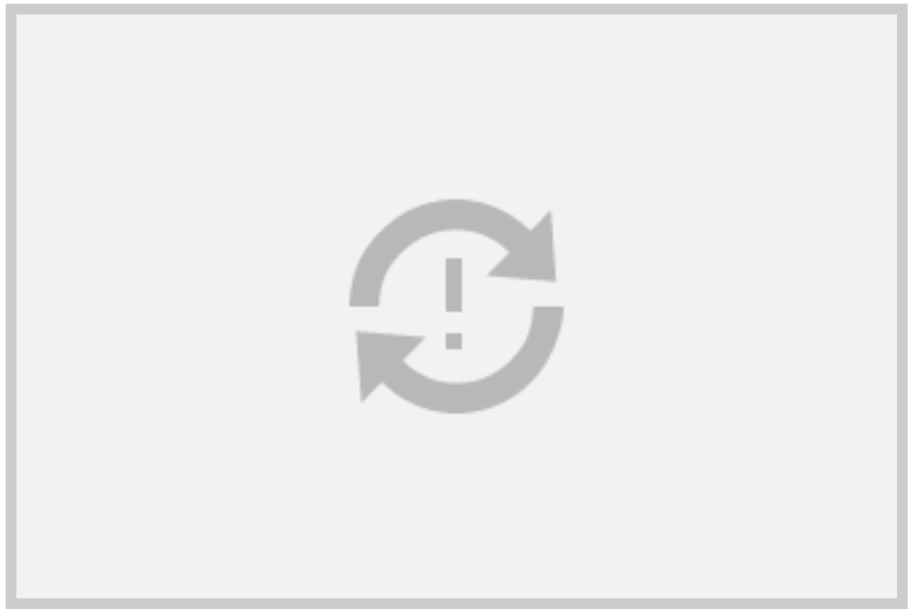


Results for check the Homoscedasticity satisfaction

Residual & fitted plot for best selected predictors



Residual & fitted plot for Log predictors



Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	10.5886	2.64715	458682.59	0.000
log(engine size)	1	0.0001	0.00007	12.69	0.000
log(hwy)	1	0.0006	0.00065	112.05	0.000
log(city)	1	0.0006	0.00058	99.81	0.000
log(comb)	1	0.0010	0.00104	180.76	0.000
Error	874	0.0050	0.00001		
Lack-of-Fit	689	0.0050	0.00001	15.38	0.000
Pure Error	185	0.0001	0.00000		
Total	878	10.5936			

Analysis Variance check the significance

$H_0 : \beta_0 = 0$
$H_1 : \beta_0 \neq 0$ } **Hypothesis**

P- Value < 0.05, So, null hypothesis can be rejected. It means our model is significant at 5% significant level.

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0024023	99.95%	99.95%	99.95%

Check the model accuracy

R^2 is 99.95%. So, the model is highly accurate.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.37206	0.00100	1368.02	0.000	
log(engine size)	0.002748	0.000771	3.56	0.000	2.85
log(hwy)	0.1769	0.0167	10.59	0.000	429.19
log(city)	0.2536	0.0254	9.99	0.000	1400.44
log(comb)	0.5660	0.0421	13.44	0.000	3258.05

Coefficients

All P – Values of predictor variables are less than 0.05. So, there is a statistically significant relationship between response and each predictor variables.

Regression Equation

log(y) = 1.37206 + 0.002748 log(engine size) + 0.1769 log(hwy) + 0.2536 log(city)
+ 0.5660 log(comb)

Regression Line

when every predictor variable remains unchanged,

If the predictors were increased, then response variable also increased.

❖ Check the functional relationship between response variable & predictor variables

Correlation: CO2 Emissions(g/km), log(engine size), log(hwy), log(city), log(comb)

	CO2 Emissions(g/	log(engine size)	log(hwy)	log(city)
log(engine size)	0.831			
	0.000			
log(hwy)	0.964	0.753		
	0.000	0.000		
log(city)	0.969	0.798	0.923	
	0.000	0.000	0.000	
log(comb)	0.986	0.798	0.967	0.990
	0.000	0.000	0.000	0.000

Cell Contents: Pearson correlation
P-Value

Correlation Coefficient Scale

+ r values	Positive	– r values	Negative
1.0	Perfect +	–1.0	Perfect –
.8 to .99	Very strong+	–.8 to –.99	Very strong–
.6 to .8	Strong +	–.6 to –.8	Strong –
.4 to .6	Moderate +	–.4 to –.6	Moderate –
.2 to .4	Weak +	–.2 to –.4	Weak –
0 to .2	Very weak +	0 to –.2	Very weak –

The response variable and all predictor variables have a very strong positive correlation, as indicated by the correlation coefficient scale above. (Since every number falls between 0.8 and 0.99)

Conclusion

For graphical analysis part

- We identified four main factors linked to CO_2 emissions in cars.
  1. Number of cylinders
  2. Fuel Type
  3. Engine size
  4. Fuel consumption (Highway/ City/ Combing rating)

- Mainly MINI, Subaru, Toyota, Ford and Honda vehicles have the least fuel consumption.
- Bentley, Mercedes – Benze and Chrysler vehicles have the most Fuel consumption under each vehicle class in all City roads, Highway roads and combined.
- Regular Gasoline can be considered as the Fuel type which emits the least amount of CO_2.

For regression analysis part

- According to regression model, we identified only engine size and fuel consumption is affect to CO_2 emissions in cars.

Suggestions to reduce CO_2 emission of vehicles

- ✔ Encouraging people to use the vehicle brands that consume a less amount of Fuel as there is a clear relationship between Fuel consumption and CO_2 Emission.

  Ex: MINI, Kia, Volvo, Ford, Honda, Toyota

- ✔ People awareness programs about the vehicle's CO2 emission and its effect on Global warming and to consume electrical vehicles, hybrid vehicles in future.

- ✔ People encourage to maintain their car (get regular tune-ups, follow the manufacturer's maintenance schedule and use the recommended motor oil)

- ✔ Also, we can suggest Drive less some facts;

1. Use public transportation modes as possible people can.
2. Promote cycle rides and walks among the people.


Any
QUESTIONS..?


THANK YOU!
Reduce
CO_2 emission
VECTOR ILLUSTRATION
CO_2
Min