Title :

“TO CHECK IMPACT OF VARIOUS COMPONETS OF AUTOMOBILE ON PRICE.”

Anekant Education Society's
Tuljaram Chaturchand College of Arts, Commerce and Science,
Baramati - 413102
Autonomous

A project report on

"TO CHECK IMPACT OF VARIOUS COMPONENTS OF AUTOMOBILE ON PRICE"

SUBMITTED TO

DEPARTMENT OF STATISTICS

T.Y.B.Sc. (Statistics)

By

Mr. Gaurav Balaso Bhoite   Mr. Makarand Sanjay Dupargude
Mr. Tejas Manik Dham   Mr. Suraj Ramchandra Jagtap
Mrs. Ankita Dattatray Kale

Under the Guidance of
Miss. Kale K.C
(2021-22)

अनेकान्त एज्युकेशन सोसायटीचे
सिद्धिरनेकान्तात्
तुळजाराम चतुरचंद महाविद्यालय,बारामती

Anekant Education Society's
Tuljaram Chaturchand College of Arts, Commerce and Science,
Baramati – 413102

DEPARTMENT OF STATISTICS

CERTIFICATE

This is to Certify that Bhoite Gaurav B., Dupargude Makarand S.,Dham Tejas M., Jagtap Suraj R., kale Ankita D., are the regular students of Department of Statistics. A project on **"Impact of Various Components of Automobile On Price"** is submitted in the partial fulfilment of the program in T.Y.B.Sc. to the Department Of Statistics, Tuljaram Chaturchand College of Arts, Commerce and Science, Baramati.

This project has been conducted under my Supervision and Guidence.

Place: Baramati
Date: 19/04/2022

Prof. Mrs. Kale K.S.		Dr. Jagtap A.S.
Project Guide	**Examiner**	**Head of Department**

TABLE OF CONTENT:

ACKNOWLEDGEMENT:

We wish to express our deep sense of gratitude to Miss. Kalyani C. Kale, Department of Statistics, Tuljaram Chaturchand College, who guided us through the project. We want to thank the entire Department of Statistics for helping us in all respects during this work.

INTRODUCTION:-

The Indian Auto industry is one of the largest in the world with an annual production.

One of the fastest growing industries in the world is the automobile industry. This automobile industries even has its influence on the Indian market. Probably automobile industries occupy a large market share in the worlds market as well as the Indian market. Nearly 18%of the total national income is being incurred from the automobile industry. From this we can estimate how important is the automobile industry has a growth rate is at the average at 10-12%.

The Automotive industry in India is one of the largest in the world and one of the fastest growing globally. India's car manufacturing industry hub Consumers are very important of the survival of the Motor Vehicle manufacturing industry

The Indian Automobile Industry manufactures over 11 million vehicles and exports about 1.5 million each year. The dominant products of the industry are two-wheeler with a market share of over 75% and passenger cars with a market share of about 16%. Commercial vehicles and three-wheeler share about 9% of the market between them.

About 91% of the vehicles sold are used by households and only about 9% for commercial purposes. The industry has a turnover of more than USD $35 billion and provides direct and indirect employment to over 13 million people.

ABSTRACT :

The Aim of this project is to find out Which factors affect the price of Cars. Now a days, most of people wishes to buy a car. The Characteristics of cars are Displacement (cc), colour, Seating Capacity, length, width and mileage.

There are many factors affecting the price of cars i.e Fuel capacity , Ground clearness, shape and No of Airbags are the major determinants of the price of cars.

So we decide to study on which factors affecting on the price of cars. We fit a regression model to predict the price of car of above Characteristics. We found that the price of car is mainly depends on Displacement(cc) and Fuel tank Capacity (litre).

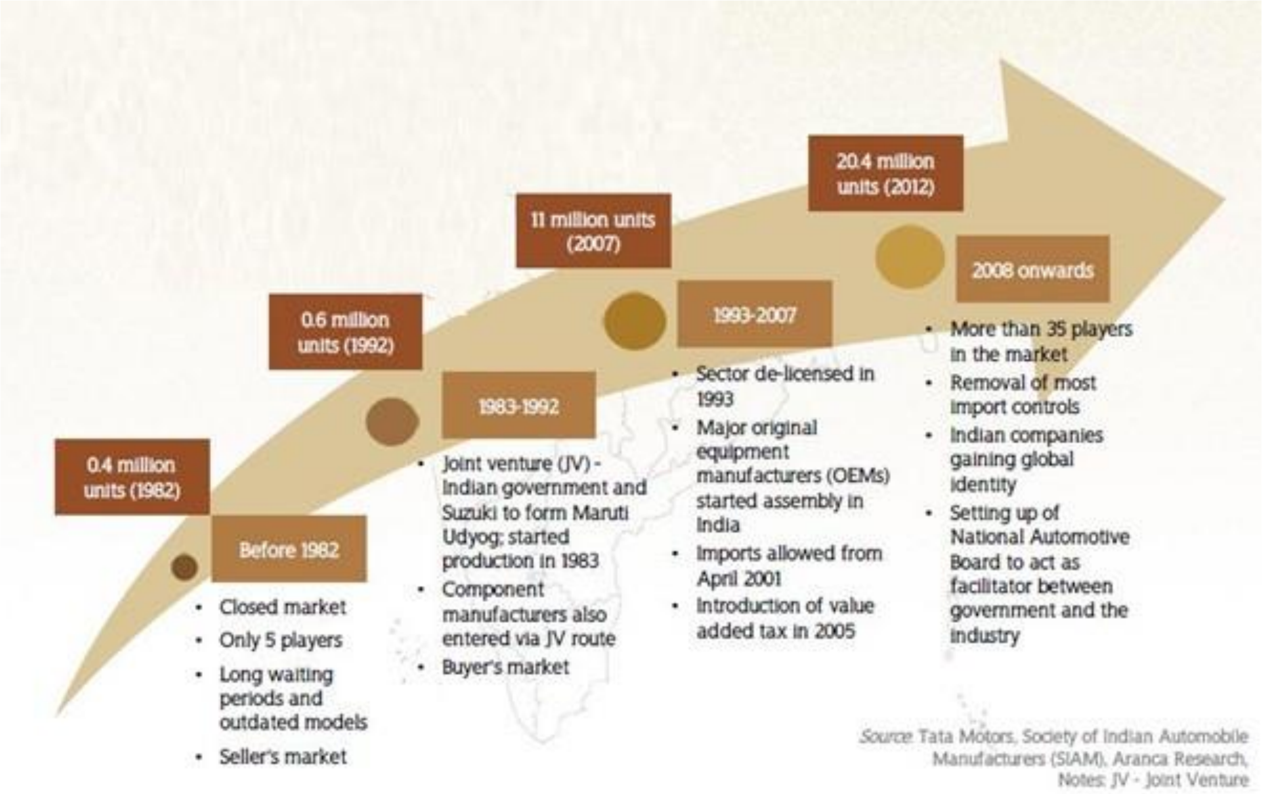
From our project we can decide to predict the best car with comfort zone and affordable.

- **KEYWORDS :** 1. Graphical Representation
  2. Correlation
  3. Testing of hypothesis
  4. Regression Analysis.

OBJECTIVES:

- To find Correlation between price and factor preferring 4-wheelers.
- To compute different Types of drive wheels impact on price.
- To predict the linear regression model for Ground clearance with Height.
- To fit preferable model for Price with different components of vehicles for prediction.
- To study and analyse the customers perception regarding the usefulness/utility of cars.

EVOLUTION OF THE INDIAN AUTOMOBILE SECTOR :-



0.4 million units (1982)

Before 1982

- Closed market
- Only 5 players
- Long waiting periods and outdated models
- Seller's market

0.6 million units (1992)

1983-1992

- Joint venture (JV) - Indian government and Suzuki to form Maruti Udyog; started production in 1983
- Component manufacturers also entered via JV route
- Buyer's market

11 million units (2007)

1993-2007

- Sector de-licensed in 1993
- Major original equipment manufacturers (OEMs) started assembly in India
- Imports allowed from April 2001
- Introduction of value added tax in 2005

20.4 million units (2012)

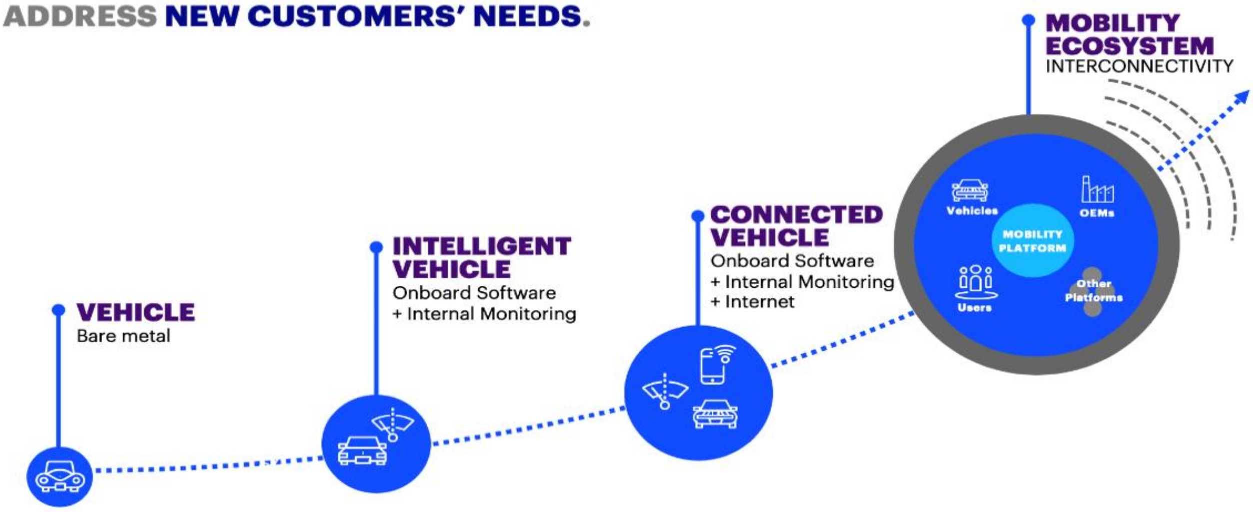
2008 onwards

- More than 35 players in the market
- Removal of most import controls
- Indian companies gaining global identity
- Setting up of National Automotive Board to act as facilitator between government and the industry

Source: Tata Motors, Society of Indian Automobile Manufacturers (SIAM), Aranca Research, Notes: JV - Joint Venture

FROM SINGLE PRODUCTS TO MOBILITY ECOSYSTEMS

THE AUTOMOTIVE INDUSTRY IS TRANSFORMING TO ADDRESS NEW CUSTOMERS' NEEDS.



MOTIVATION :

In our curriculum, there is a course "REGRESSION ANALYSIS and PARAMETRIC TESTS" which applicable more in forecasting and prediction And also for check the dependency and Independency between two Attributes.

We have experience of handling different Data sets in practical during our TY.BSC. During learning regression also get Theoretical knowledge but we where much interested to knowing how data is analysed in Actual sense.

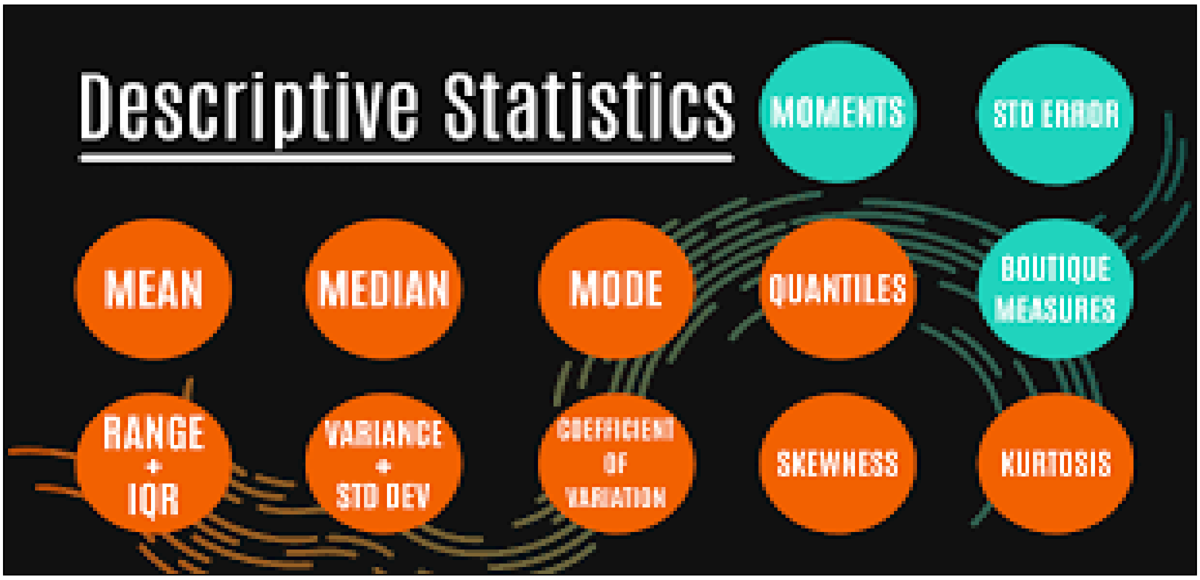
So, we decided to carry out "Regression Analysis and Parametric Tests" in our Project.

METHODOLOGY :

As we decided to work on Automotive Industry. We decided to Study Different Brands of Vehicles of India. All information used in this project is collected from online source and various websites. Taking this data as a basis of research we decided to work on different components of a car that will give us option to choose the perfect, comfortable and luxurious car. And will be developed so as to achieve the maximum output from the prototype engine and to achieve the longest mileage for many cars in future.

In this project the data contains different components and parts of vehicles of various brands in India.

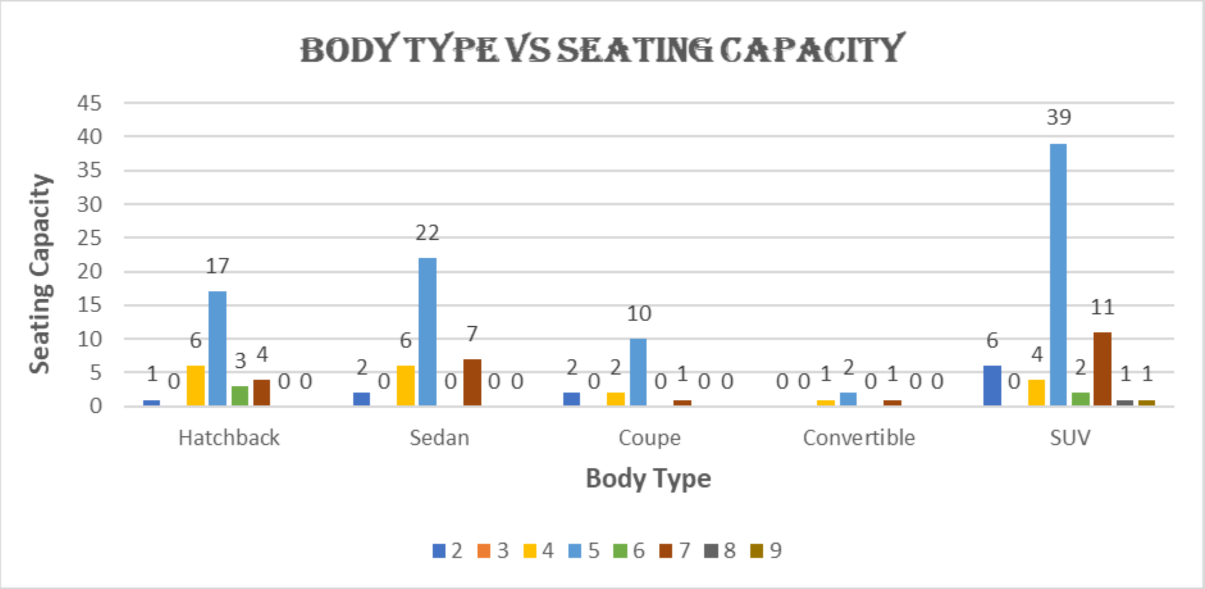
Firstly we drew a scatter plot and box-plot with fitted lines to find out the correlation between price and other different components of vehicles. Then we developed models for prediction of the vehicles price which will give customer a choice of budget.For such analysis we used some software like R-software, MS-Excel, Tableau Software etc.


Descriptive Statistics
MOMENTS
STD ERROR
MEAN
MEDIAN
MODE
QUANTILES
BOUTIQUE MEASURES
RANGE + IQR
VARIANCE + STD DEV
COEFFICIENT OF VARIATION
SKEWNESS
KURTOSIS

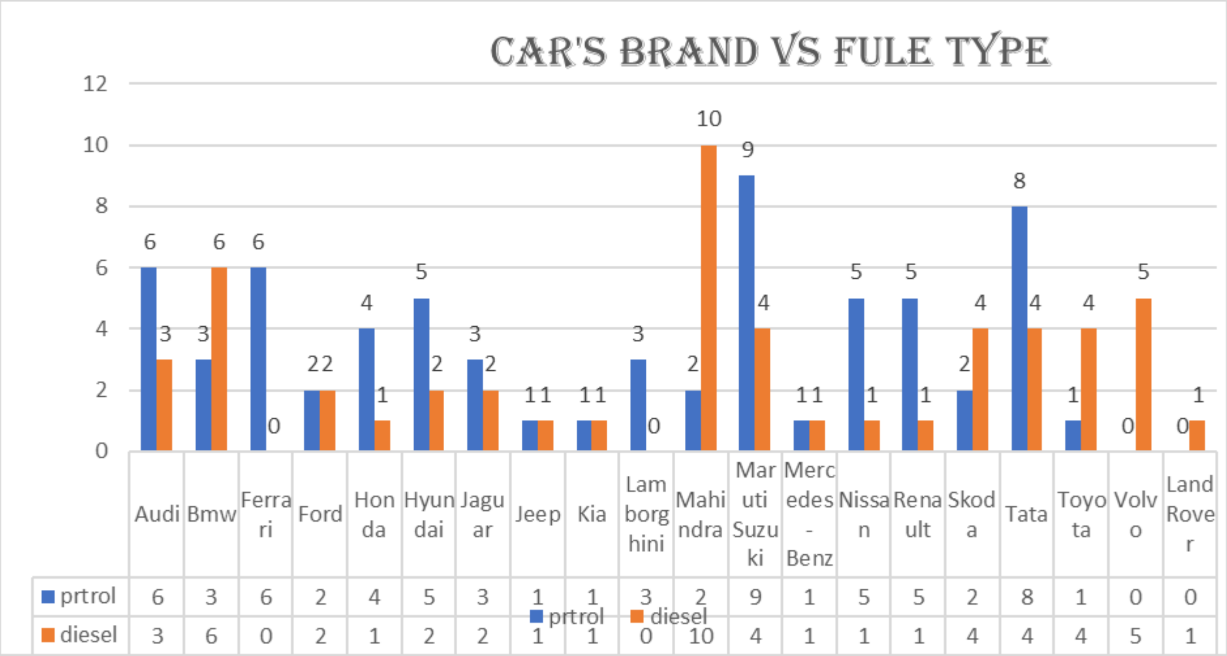
	Price (Rs)	*Displacement(cc)*	*Gears*	*Number of Airbags*
Mean	9316422.982	2376.017751	6.24852071	4.467455621
Standard Error	1508724.499	111.6241595	0.102029651	0.215952206
Median	2371858	1984	6	4
Mode	23290000	1498	5	2
Standard Deviation	19613418.49	1451.114073	1.326385462	2.807378677
Sample Variance	3.84686E+14	2105732.053	1.759298394	7.881375035
Kurtosis	46.59342834	2.049104698	-0.937859607	-0.647374277
Skewness	5.771223375	1.554824526	0.493162149	0.563246247
Range	191906490	7369	6	13
Minimum	236447	624	4	1
Maximum	192142937	7993	10	14
Sum	1574475484	401547	1056	755
Count	169	169	169	169
C.I. (95%)	2978501.564	220.3667626	0.201425426	0.426329647

	Fuel Tank Capacity (Litres)	**City Mileage (km/litre)**	**Highway Mileage (km/litre)**
Mean	58.77159763	14.36035503	18.71035928
Standard Error	1.413167703	0.590099631	1.365925906
Median	57	13.5	16.8
Mode	60	18	15.4
Standard Deviation	18.37118015	7.671295204	17.65165284
Sample Variance	337.5002599	58.84877011	311.5808481
Kurtosis	-0.614733338	30.92800831	81.25165226
Skewness	0.450572202	3.802953126	8.29373975
Range	81	78	201
Minimum	24	2	4
Maximum	105	80	205
Sum	9932.4	2426.9	3124.63
Count	169	169	167
Confidence Level (95.0%)	2.789854753	1.164965953	2.696826417

Exploratory Data Analysis :-

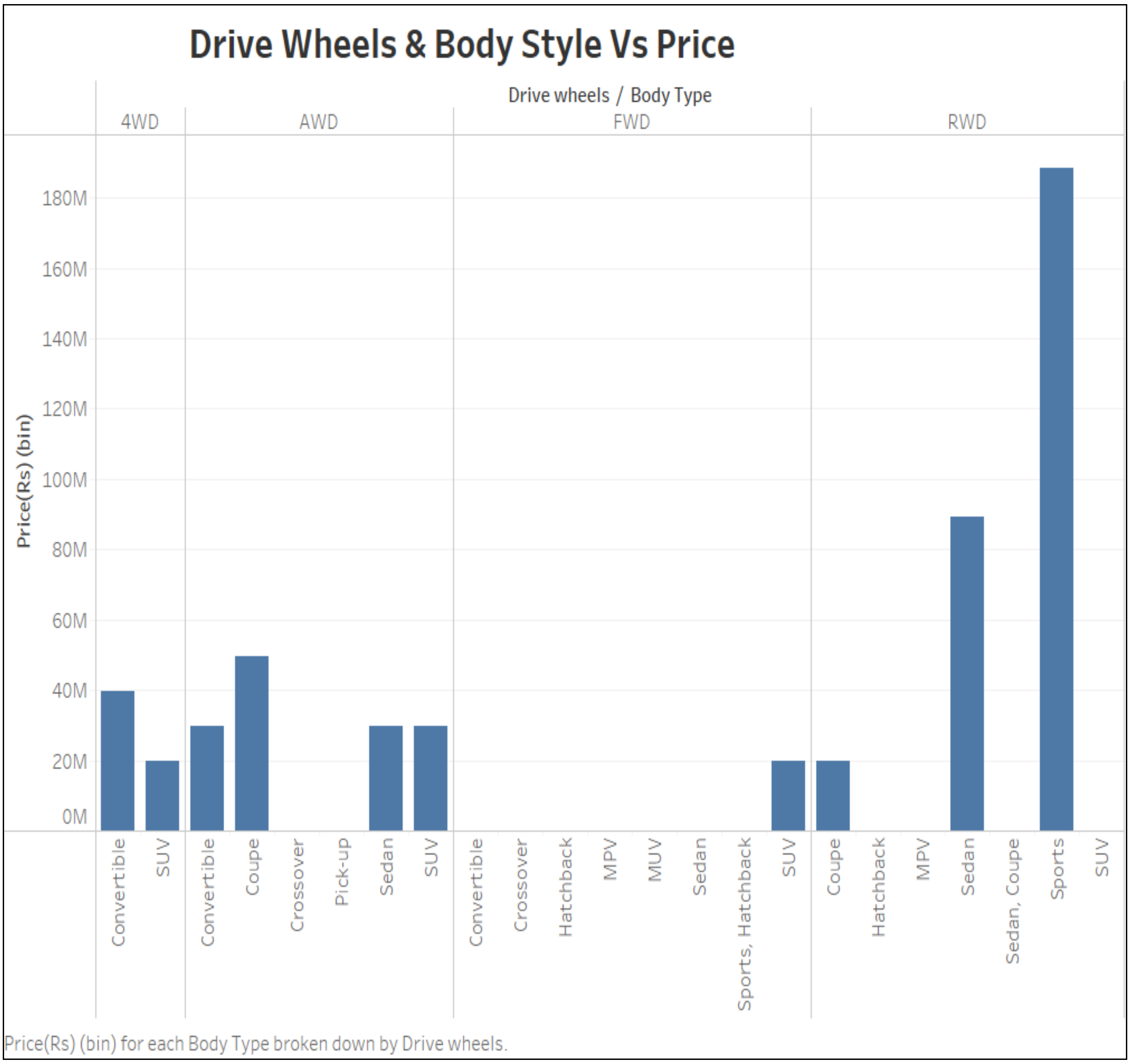


Conclusion : Here, We see that the distribution of price between the different body-style categories have a significant overlap, and so body-style would not be a good predicator of price.



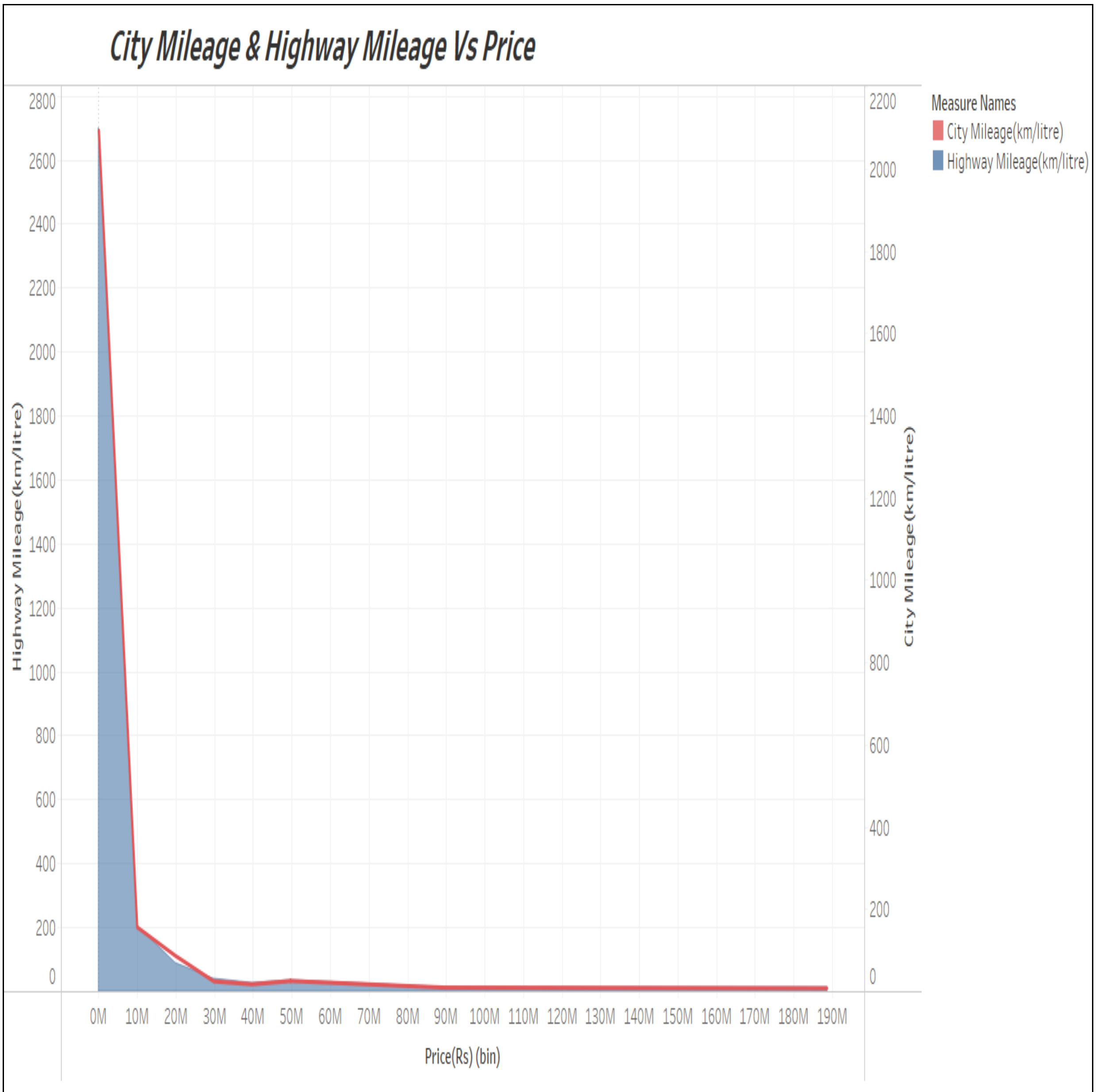
	Audi	Bmw	Ferrari	Ford	Honda	Hyundai	Jaguar	Jeep	Kia	Lamborghini	Mahindra	Maruti Suzuki	Mercedes - Benz	Nissan	Renault	Skoda	Tata	Toyota	Volvo	Land Rover
prtrol	6	3	6	2	4	5	3	1	1	3	2	9	1	5	5	2	8	1	0	0
diesel	3	6	0	2	1	2	2	1	1	0	10	4	1	1	1	4	4	4	5	1

Conclusion: - Here, we observed that From All observations there is rare difference in uses of Petrol and Diesel Cars.



Conclusion:-

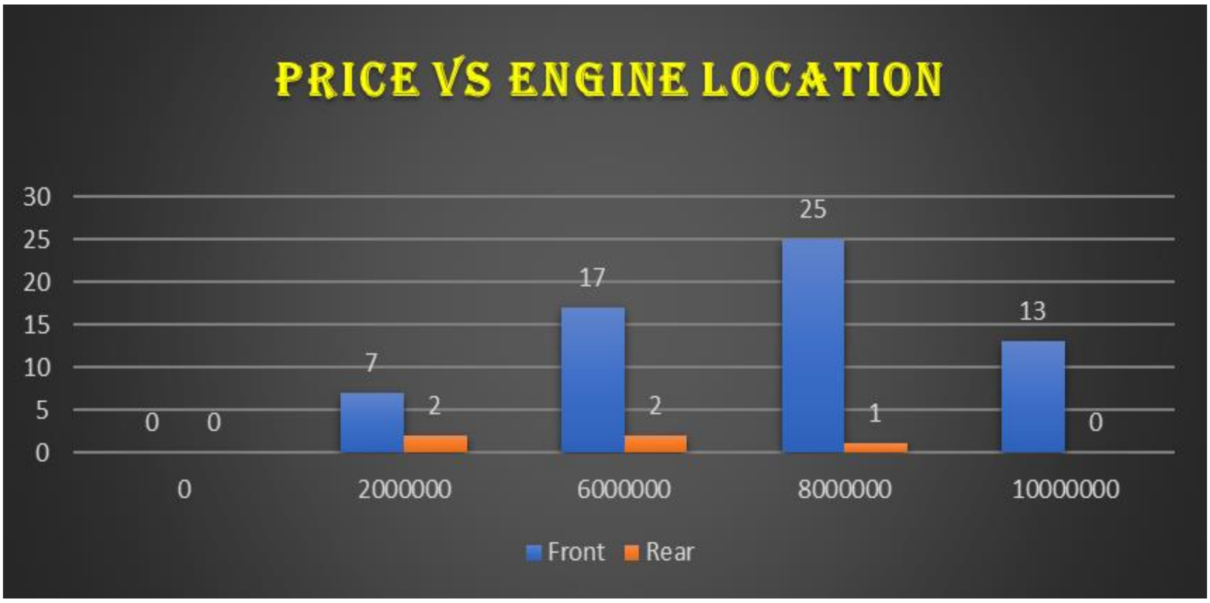
In this plot the target variable (price) proportional to bar graph with respect to the variables 'Drive wheels' and 'Body style' in the horizontal axis respectively.this allows us to visualize how the price is realted to 'Drive wheels' and 'Body style'



The trends of Highway Mileage(km/litre) and City Mileage(km/litre) for Price(Rs) (bin). Color shows details about Highway Mileage(km/litre) and City Mileage(km/litre).

Conclusion:

This graph Visualised the trends of Highway Mileage and City Mileage for Price by using Dual axis command.



Conclusion :

Here, we observed that the distribution of price between these two engine-location categories, front &rear are distinct enough to take engine-location as a potential good predicator of price.

STATISTICS
DATA ANALYSIS

• TESTING OF HYPOTHESIS:

1.Chi-Squared Test For Independence Of Attributes: -

In this test, we want to test the

Null Hypothesis H_0 : The Two Attributes A and B are Independent

Against

Alternative H_1 : The Two Attributes A and B are Dependent.

The R-software Following Command is used for performing the test: -

```
chisq.test(y,conf.level=correct=F)
```

- ❖ Let’s we have to calculate which components of vehicles depends on Price (Rs).

  By Using R-Software-

- **Price (Rs) and Displacement (cc):**

```
x=scan("clipboard")
y=scan("clipboard")
z=c(x,y)
mx=matrix(z,nrow=2,ncol=169)
chisq.test(mx,correct=T)

Pearson's Chi-squared test
data:  mx
X-squared = 328933676, df = 168, p-value < 2.2e-16
```

Conclusion: -

Here P-value (2.2e-16) is
Less Than Level of Significance
Alpha = 0.05. Hence, we reject the
Null hypothesis.

Therefore, the price of vehicle is dependent on displacement(cc).

- **Price (Rs) and No. of Airbags: -**

```
y=scan("clipboard")
 x=scan("clipboard")
 z=c(x,y)
 mx=matrix(z,nrow=2,ncol=169)
 chisq.test(mx,correct=F)

        Pearson's Chi-squared test
data:  mx
X-squared = 345185651, df = 168, p-value < 2.2e-16
```

Conclusion: -

Here P-value (2.2e-16) is
Less Than Level of Significance
Alpha = 0.05. Hence, we reject the
Null hypothesis.

Therefore, price is dependent on no of air bags present in the cars.

- **Price (Rs) and No. of Gears: -**

```
y=scan("clipboard")
 x=scan("clipboard")
 z=c(x,y)
 mx=matrix(z,nrow=2,ncol=169)
 chisq.test(mx,correct=F)

        Pearson's Chi-squared test
data:  mx
X-squared = 345185609, df = 168, p-value < 2.2e-16
```

Conclusion: -

Here P-value (2.2e-16) is
Less Than Level of Significance
Alpha = 0.05. Hence, we reject the
Null hypothesis.

Therefore, The price is dependent on no of gears present in the car.

- **Price (Rs) and Fuel Tank Capacity (litres): -**

```
 y=scan("clipboard")
 x=scan("clipboard")
 z=c(x,y)
 mx=matrix(z,nrow=2,ncol=169)
 chisq.test(mx,correct=F)

        Pearson's Chi-squared test
data:  mx
X-squared = 345185609, df = 168, p-value < 2.2e-16
```

Conclusion: -

Here P-value (2.2e-16) is
Less Than Level of Significance
Alpha = 0.05. Hence, we reject the
Null hypothesis.

Therefore, the price is dependent on Fuel Tank Capacity (Litres).

2. Paired t-test: -

Paired t-test is used when two samples of same size and which are not independent but correlated.

The R-software Following Command is used for performing the test: -

```
t.test(x,y,paired=T,conf.level=c)
```

❒ Let's we have to calculate the correlation between City Mileage (per km's) with respect to Highway Mileage (per km's).

- **City Mileage(km/litre) And Highway Mileage (km/litre): -**

```
> x=scan("clipboard")
> y=scan("clipboard")
> t.test(x,y,paired=T)

        Paired t-test
data:  x and y
t = -3.9524, df = 168, p-value = 0.0001138
alternative hypothesis:
true difference in means is not equal to 0
95 percent confidence interval:
 -6.660114 -2.223081
sample estimates:
mean of the differences
             -4.441598
```

Conclusion: - Here P-value (0.0001138) is Less Than Level of Significance Alpha = 0.05. Hence, we reject the

Null hypothesis Also we observed that 95% C.I. (-6.660114, -2.223081)

Hence, we Conclude that there is change between **"City Mileage(km/litre) And Highway Mileage (km/litre)".**

3.ANOVA (one way classification):

The Analysis of Variance (ANOVA) is a statistical method used to test whether there are significant differences between the means of two or more groups. ANOVA returns two parameters:

P-value: P-value tells how statistically significant is our calculated score value.

If our price variable is strongly correlated with the variable we are analysing, expect ANOVA to return a sizeable F-test score and a small p-value.

•Drive Wheels: -

Since ANOVA analyses the difference between different groups of the same variable, the group by function will come in handy. Because the ANOVA algorithm averages the data automatically, we do not need to take the average before hand.

❒ Let's see if different types drive-wheels impact on price.

By using R-Command: -

```
> x1=scan("clipboard")
> x2=scan("clipboard")
> d=stack(list(b1=x1,b2=x2))
> names(d)
[1] "values" "ind"
> av1=aov(values~ind,data=d)
> summary(av1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ind	1	7.334e+15	7.334e+15	38.13	1.91e-09
Residuals	336	6.463e+16	1.923e+14		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Conclusion: -

If P value(given by Pr(>F)) is less than level of significance, then we reject corresponding null hypothesis.

Hence, we conclude that there is impact of Drive Wheels on price.

4.MODEL DEVELOPMENT: -

- Objectives: -

Develop prediction models –

In this section, we will develop several models that will predict the price of the car using the variables or features. This is just an estimate but should give us an objective idea of how much the car should cost.

Some questions we want to ask in this module -

- do I know if the dealer is offering fair value for my trade-in?
- do I know if I put a fair value on my car?
- Data Analytics, we often use Model Development to help us predict future observations from the data we have.

❒ A Model will help us understand the exact relationship between different variables and how these variables are used to predict the result.

4.1 Simple Linear Regression Model: -

Simple Linear Regression is a method to help us understand the relationship between two variables:

- The predictor/independent variable (X)
- The Dependent variable (that we want to predict) (Y)

The result of Linear Regression is a linear function that predicts the response (dependent) variable as a function of the predictor (independent) variable.

Y: Dependent Variable (Response)

X: Independent Variable (Predictor)

Linear Function:

Y hat = a + bX

- a refers to the intercept of the regression line 0, in other words: the value of Y when X is 0
- b refers to the slope of the regression line, in other words: the value with which Y changes when X increases by 1 unit.

❒ We will create a linear function with “Ground Clearance” as the dependent variable and “Height” as a independent variable.

By using R-Command: -

```
> y=scan("clipboard")
> x=scan("clipboard")
> a1=lm(y~x)
> summary(a1)
```

```
Call:
lm(formula = y ~ x)
```

Residuals:

Min	1Q	Median	3Q	Max
-100.62	-14.26	0.27	15.91	320.18

Coefficients:

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	18.97032	23.04069	0.823	0.411
x	0.09800	0.01445	6.780	1.98e-10

S. codes: 0 ‘* * *’ 0.001 ‘* *’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 36.73 on 167 degrees of freedom

Multiple R-squared: 0.2159, Adjusted R-squared: 0.2112

F-statistic: 45.97 on 1 and 167 DF, p-value: 1.977e-10

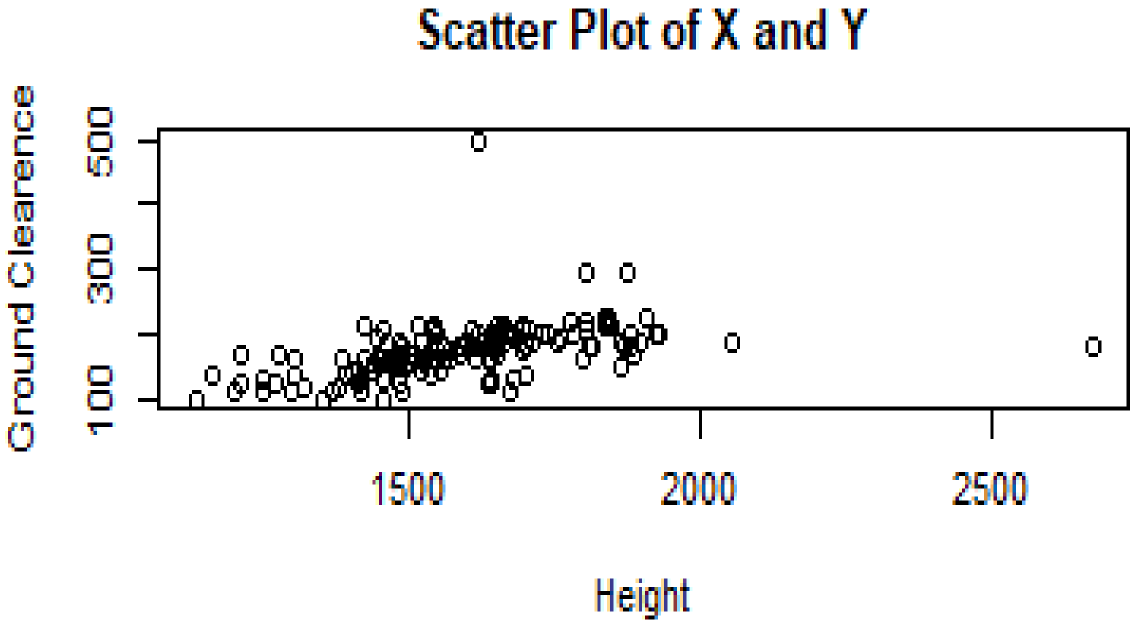
```
> cor(x,y)

[1] 0.4646031
```

```
>plot(x,y,xlab="Height",ylab="Ground Clearence",main="Scatter Plot of X and Y")
> abline(a1)
>text(98,1200,"Line of best Fit Y=18.97032+0.09800X")
```

- We should get the final Estimated Linear Model with actual values:

Ground Clearance = 18.97032 + 0.09800* Height

❒ **Let's visualize height as a potential, Predictor variable of Ground Clearance: -**



- **Interpretation: -**

We can see from this plot that Ground Clearance is positively correlated to height, since the regression slope is Positive. One thing to keep in mind when looking at a regression plot is to pay attention to how Close the data points are around the regression line. This will give you a good indication of the variance of the data, and whether a linear model would be the best fit.

4.2 Multiple Linear Regression Model: -

If we want to use more variables in our model to predict car price, we can use Multiple Linear Regression. Multiple Linear Regression is very similar to Simple Linear Regression, but this method is used to explain the relationship between one continuous response (dependent) variable and two or more predictor (independent) variables. Most of the real-world regression models involve multiple predictors. We will illustrate the structure by using four predictor variables, but these results can generalize to any integer:

Y: Response Variable
X_1: Predictor Variable 1st
X_2: Predictor Variable 2nd
X_3: Predictor Variable 3rd
X_4: Predictor Variable 4th

The Equation is Given By-

$$\hat{Y} = a+b_1X_1+b_2X_2+b_3X_3+b_4X_4$$

- From this Section We know that the good predictors of price could be: -

★ **Displacement**
★ **No. of Airbags**
★ **No. of Gears**
★ **Fuel Tank Capacity**

❒ Let's develop a model using this variable as a predictor variables: -

By using R-Command: -

```
> y=scan("clipboard")
> x1=scan("clipboard")
> x2=scan("clipboard")
> x3=scan("clipboard")
> x4=scan("clipboard")
> a1=lm(y~x1+x2+x3+x4)
> print(a1)
```

Call:
lm(formula = y ~ x1 + x2 + x3 + x4)

Coefficients:

(Intercept)	x1	x2	x3	x4
-6764819	13378	-350493	-87593	-223294

```
> summary(a1)
```

Call:
lm(formula = y ~ x1 + x2 + x3 + x4)

Residuals:

Min	1Q	Median	3Q	Max
-36130080	-3121094	769648	3146519	117287152

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6764819	4928932	-1.372	0.1718
x1	13378	1220	10.968	<2e-16
x2	-350493	857992	-0.409	0.6834
x3	-87593	335450	-0.261	0.7943
x4	-223294	102123	-2.187	0.0302

```
Signif. codes:  0 ‘* * *’ 0.001 ‘* *’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11790000 on 164 degrees of freedom
Multiple R-squared:  0.6471,    Adjusted R-squared:  0.6385
F-statistic: 75.18 on 4 and 164 DF, p-value: < 2.2e-16
```

```
> m=cbind(y,x1,x2,x3,x4)
> pairs(m)
> cor(m)
```

	y	x1	x2	x3	x4
y	1.0000000				
x1	0.7957138	1.0000000			
x2	0.3583941	0.5147404	1.0000000		
x3	0.1942601	0.2566757	0.1428553	1.0000000	
x4	0.6193513	0.8541027	0.6004574	0.2097326	1.0000000

Conclusion :

We conclude that,64.7% of total variation in the fitted model explained by the response variable **price** and independent variables **Displacement, No of Airbags, No. of Gears, Fuel tank capacity.**

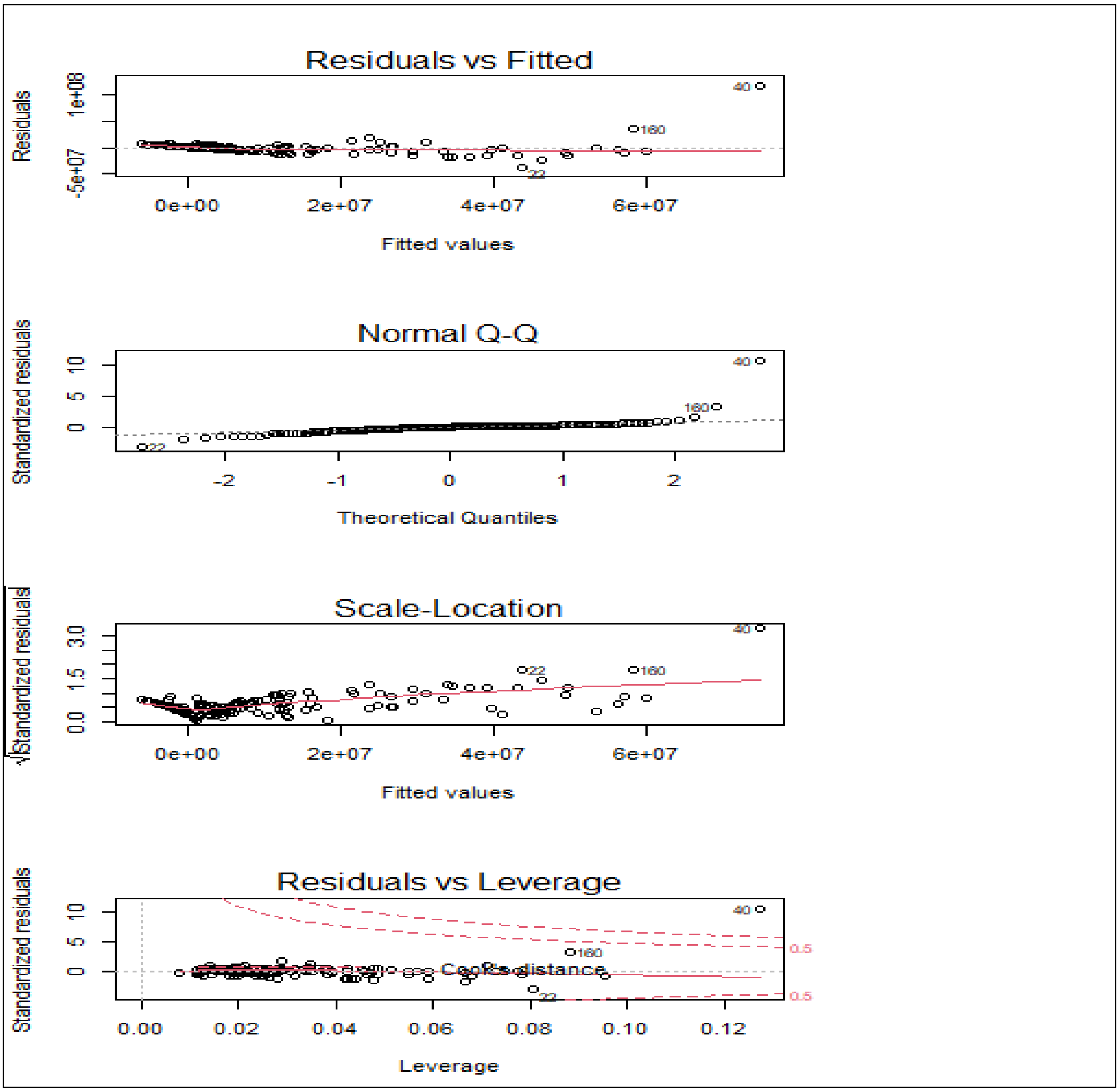
- From above R-square best fitted model for our data is ;

$$\mathbf{Y = -6764819 + 13378*X_1 - 350493*X_2 - 87593*X_3 - 223294*X_4}$$

➢ Normality Assumptions for Fitted Model :-

By using R-Command: -

```
lm(formula = y ~ x1 + x2 + x3 + x4)
> par(mfrow=c(4,4))
> plot(a1)
```



Conclusion:

Here, residual plot shouldn't show any pattern ,so error variance is constant. from QQ plot ,we can see a straight line ,so errors are normally distributed. thus both assumptions are satisfied.

4.3 Backword Elimination Method : -

The all possible regressors approach considers all possible subsets of the pool of explanatory variables and finds the model that best fits the data according to some criteria (e.g . Adjusted R^2, AIC and BIC)

These criteria assign scores to each model and choose the model with the best score.

We will use a process called backward elimination to help decide which predictors to keep in our model and which to exclude. In backward elimination, we start with all possible predictors and then use lm() to compute the model. We use the summary() function to find each predictor’s significance level.

- Here we develop a Backward Elimination model for Response Variable Price with respect to predictor Variables :

⋆ **Displacement**

⋆ **No. of Airbags**

⋆ **No. of Gears**

⋆ **Fuel Tank Capacity**

❒ Let’s develop a Backward Elimination model using this variable as a predictor variables: -

By using R-Command: -

```
>data=read.csv("C:/Users/Makarand/Desktop/stepwise.csv",
header=T)

> x=lm(Price.Rs.~ . , data=data)
> summary(x)

> step(x, direction = "backward")
```

```
Start:  AIC=5508.57

     Price. ~ Displacement.cc. + Gears + Number of Airbags +
Fuel Tank Capacity
```

	Df	Sum of Sq.	RSS	AIC
No.of Airbags	1	9.4821e+12	2.2816e+16	5506.6
Gears	1	2.3207e+13	2.2830e+16	5506.7
F.T.C..Litres.	1	6.6486e+14	2.3472e+16	5511.4
Displacement.cc.	1	1.6728e+16	3.9535e+16	5599.5

```
Step:  AIC=5506.64

    Price ~ Displacement +Gears + Fuel Tank Capacity
```

	Df	Sum of Sq.	RSS	AIC
Gears	1	2.3850e+13	2.2840e+16	5504.8
F.T.C.Litres.	1	6.6127e+14	2.3477e+16	5509.5
Disp.cc.	1	1.7002e+16	3.9818e+16	5598.8

Step: AIC=5504.82

Price. ~ Displacement. + Fuel Tank Capacity.

	Df	Sum of Sq.	RSS	AIC
F.T.C.Litres.	1	8.6783e+14	2.3708e+16	5509.1
Disp.cc.	1	1.6997e+16	3.9837e+16	5596.8

Call:
lm(formula=Price~ Displacement + Fuel Tank Capacity
data = data)

Coefficients:

(Intercept)	Displacement	Fuel Tank Capacity
-8368926	13327	-237868

Conclusion:

From Above table we observe that the model with Regressors X1 and X4 have maximum AIC value.

Here we prefer model with regressor X1 and X4 for the prediction

- So the model suggested by the Backward Elimination method is as follows:

Price = - 8368926 + 13327 * Displacement - 237868 * Fuel Tank Capacity.

CONCLUSIONS:

1. we conclude that, by using Chi-Squared Test For Independence Of Attributes that is price is Dependent Variable on Various Components of car.

2. We conclude that ,by using paired t - test for correlation between two independent Variables, and we analysed that there is change in **Mileage(km/litre)** between "**City And Highway**".

3. we conclude that there is impact of Drive Wheels on price by using ANOVA.

4. We Predict the Multiple Linear Regression model for, 64.7% of total variation in the fitted model explained by the response variable **price** and independent variables **Displacement, No of Airbags, No. of Gears, Fuel tank capacity.**

5. We prefer Model of Backward Elimination method with Regressor **Displacement** and **Fuel tank capacity** for the prediction of response variable **price .**

SCOPE:

There is no doubt that this is the most flourishing technology in automotive sector. Almost the conventional vehicles have been replaced by hybrid and many of this has smart engine driving system. As per the rising need it is sure that in future the conventional carbureted vehicles shall be replacement by this smart system. India still lag in this section. Hence there is huge scope to this.

LIMITATION:

- Unable to work on the all different components if automobile sector.
- We have study on 170 samples so we can increase sample size and Interpret again.
- Lack of previous research studies on this topic.

REFERENCES:

❖ Statistical Methods & Use of R – Software

Prof P.G. Dixit, Prof P. S. Kapare

❖ Introduction to Linear Regression Analysis

Prof Douglas C. Montgomery,

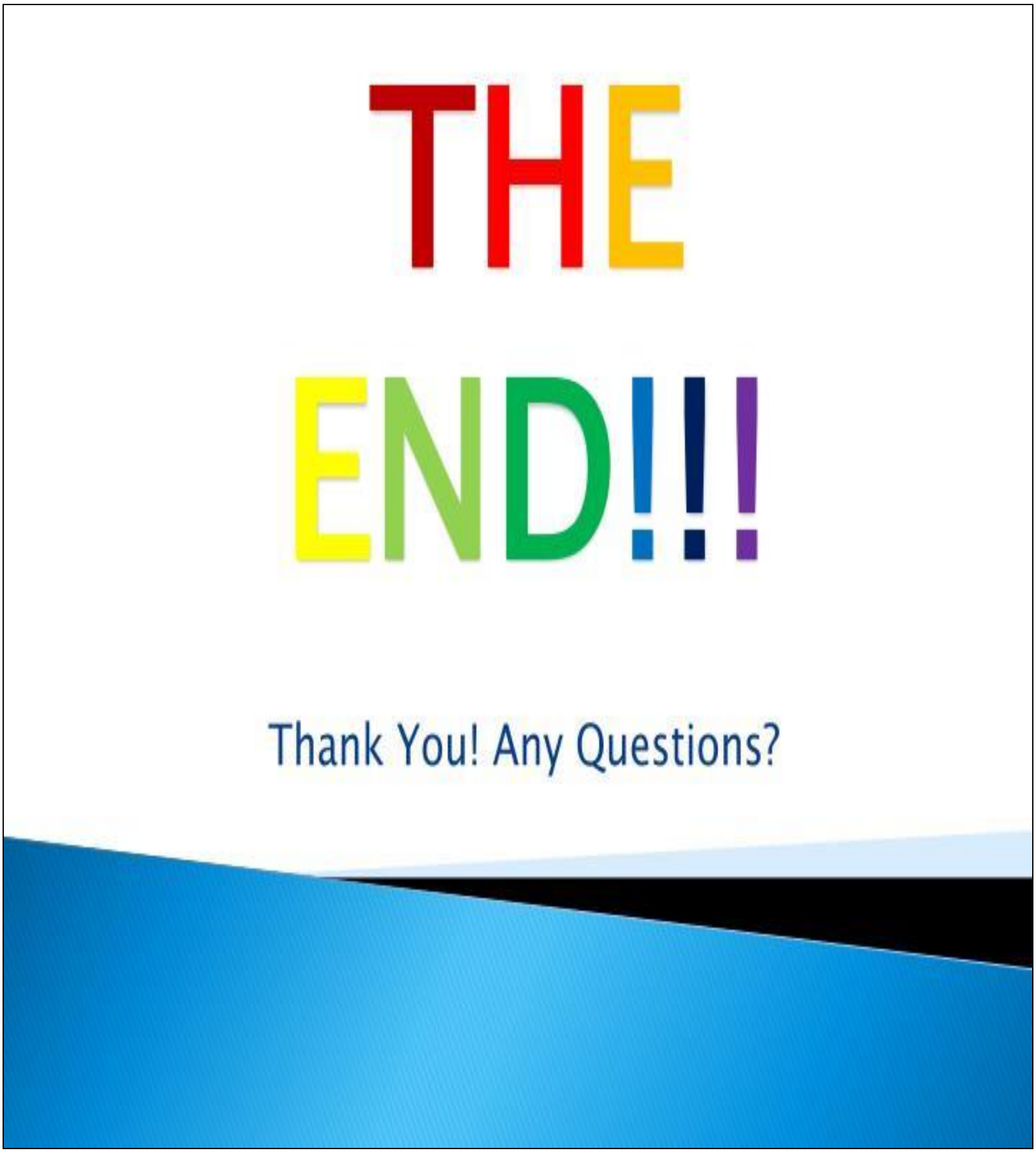
Prof Elizabeth A. Peck

❖ Statistical Computing Using R- Software

Prof Vishwas R. Pawgi

❖ Web Search-

- Kaggle website
  https://www.kaggle.com/Automobile
- Education Wikipedia, Creative Education.



THE END!!!

Thank You! Any Questions?