Sankepally Vikram Reddy
vikramreddy0307@gmail.com
F:907,Bhavyas srinivasam,opp-Nizampet
Hyderabad-500085

Overview

The project is about data collected from the sensors inserted in the car about the rpm ,speed and the fuel economy.The data is collected from Honda city iVtec S car of 1497 cc.

Goals

1. Draw the conclusions from the collected data about the idle RPM,fuel economy etc
2. Find the patterns of RPM and Speed in which the car has not given the economy range.

Specifications

The Honda City i-VTEC S is the base petrol variant in the City range and priced at Rs 8.72 lakh (ex-showroom, Delhi). It comes with exterior features like front Chrome grill, moulding line, dual-barrel headlamps with integrated LED DRLs, wraparound tail lamps, rear Chrome garnish, 15-inch Steel wheels

• Engine Type:	i VTEC Engine
• Engine Displacement:	1497 cc
• Fuel Type:	Petrol
• Power:	117.3bhp@6600rpm
• Torque:	145Nm@4600rpm
• No Of Cylinders:	4
• Transmission:	Manual
• Gear Box:	5 Speed
• Drive Type:	FWD
• Paddle Shift:	N
• Kerb Weight:	1058kg

Contents:

Chapter 4:

Code

- R

Please go through page number 10 and 11 to understand the flow of analysis [Very IMP]

Suggestions

Chapter 1

Overview of Data:

date	**time**	**speed**
Length:1443	Length:1443	Length:1443
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character
fuel_economy	**rpm**	**coolant_temp**
Length:1443	Length:1443	Length:1443
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Variables in data:

Classes 'tbl_df', 'tbl' and 'data.frame': **1443 obs. of 6 variables:**

$ date : chr "2018/04/30" "2018/04/30" "2018/04/30" "2018/04/30" ...

$ time : chr "00:01:21" "00:03:21" "00:05:21" "00:07:21" ...

$ speed : chr "0" "0" "0" "0" ...

$ fuel_economy : chr NA NA NA NA ...

$ rpm : chr "0" "0" "0" "0" ...

$ coolant_temp : chr NA NA NA NA ...

Chapter 2:

Data Preprocessing

Converting to bins.

As it very difficult for the garage owner or car company owner to view the exact figures of the data collected and draw conclusions from it.It is very useful if we bin the ranges of the RPM and speed so that it will be easy and helpful for the person viewing it.

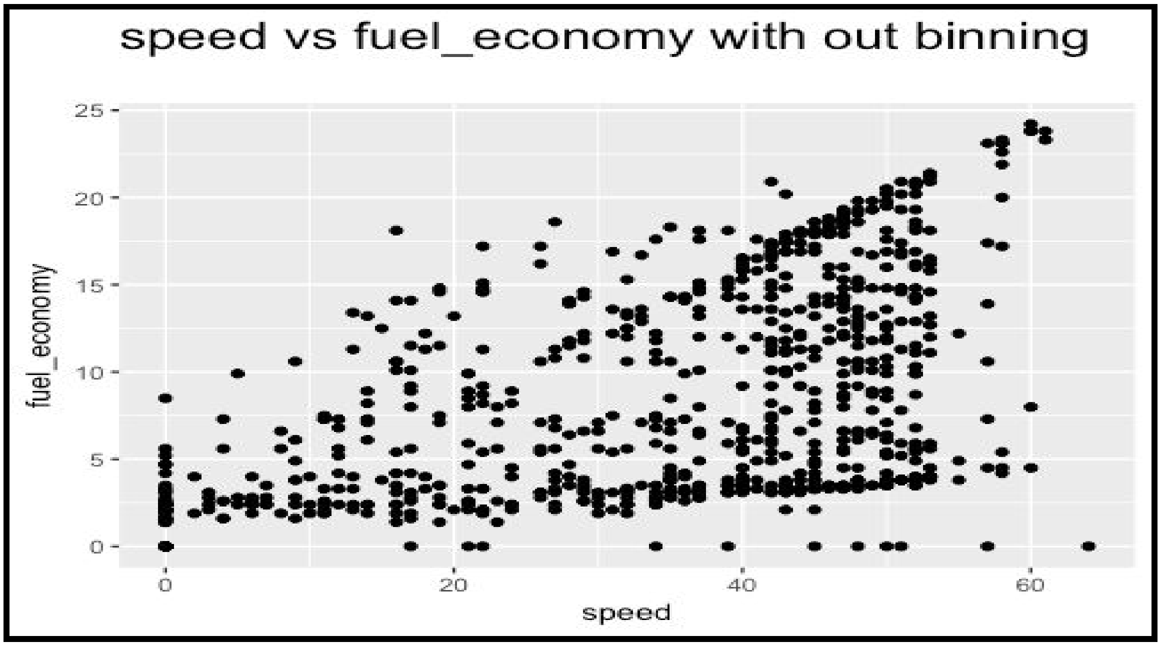
Reason

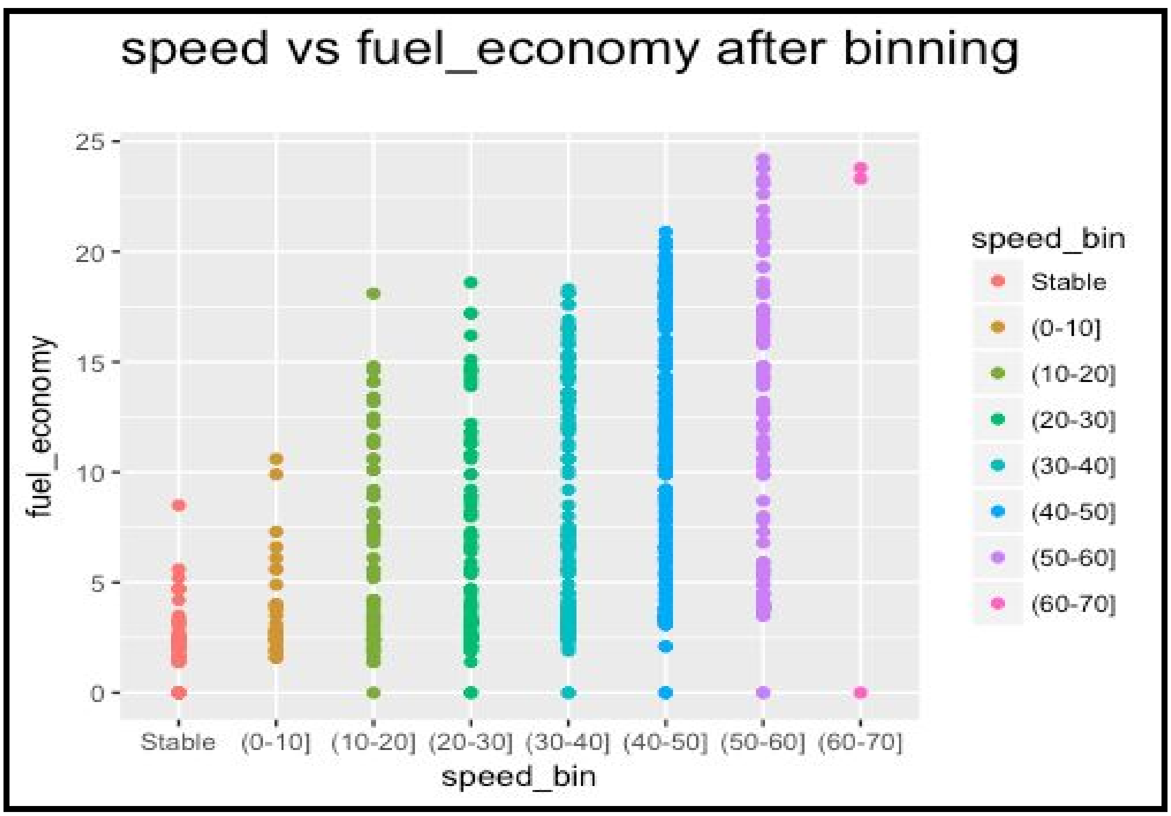
- Easy to visualize
- Helpful to draw conclusions
- Find the underlying hidden patterns in the data

Speed binning

As the speed is ranging from

Min.	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max.**
0.00	22.00	40.00	34.08	47.00	64.00



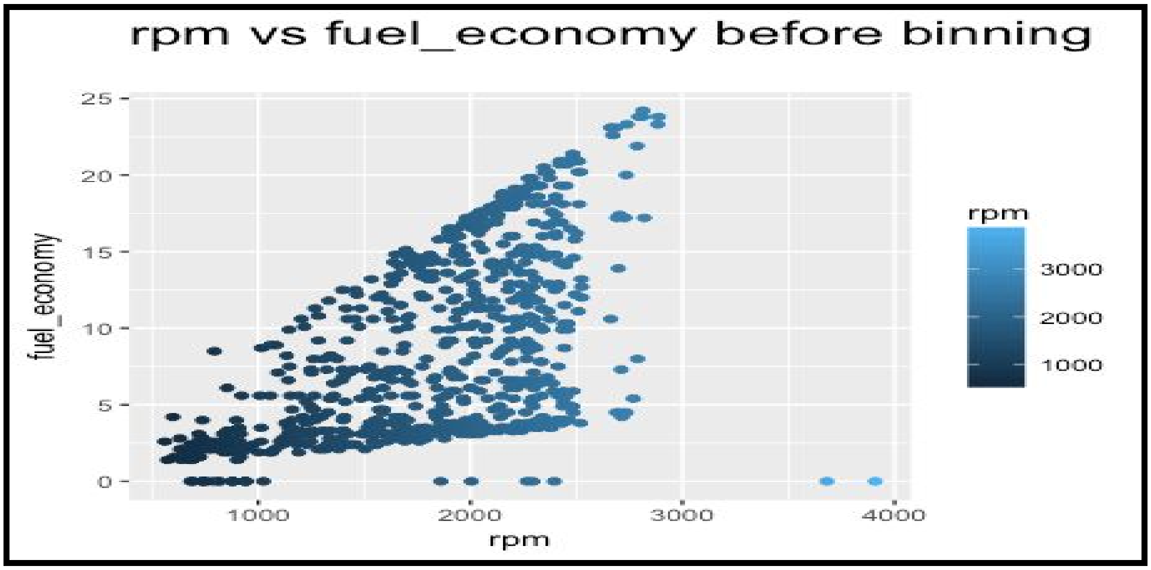


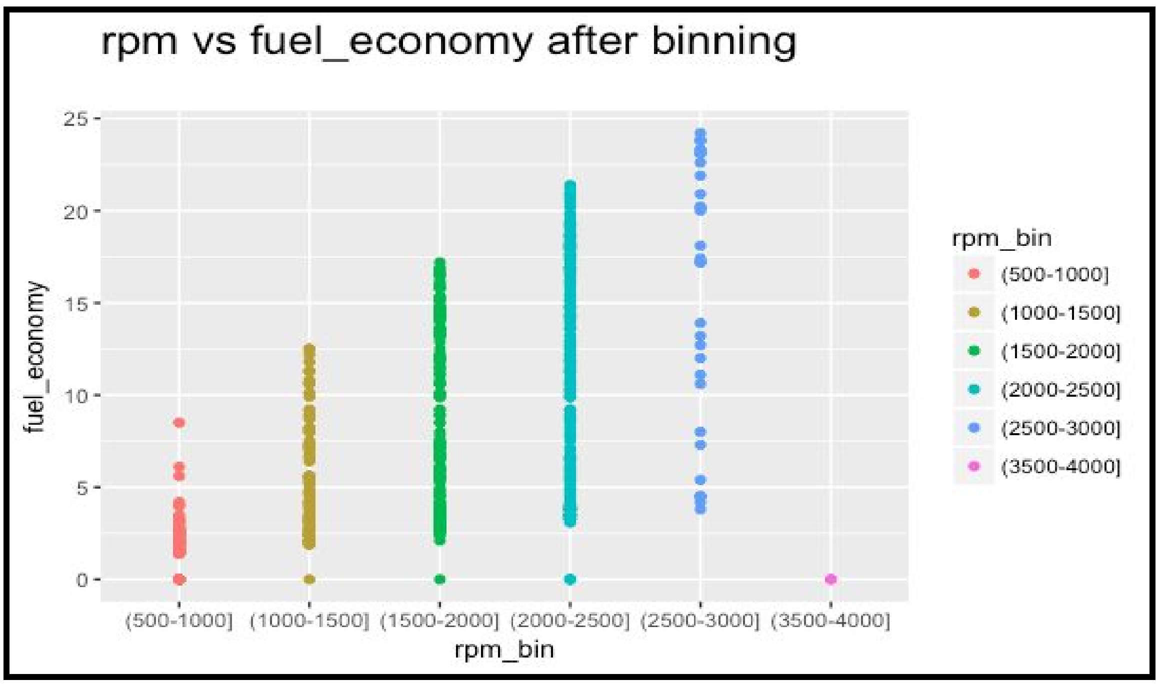
From the above graph it is very clear that after binning it is easy to draw conclusions about what is the range of fuel_economy for a certain range of speed.

Binning of RPM:

RPM of the car is of the range

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
561	1415	1927	1796	2236	3909





Extracting hour from Time.

The data include the time of the values collected from the car.So it is necessary to consider the time factor of the car to analyse deeply the performance of the car.

I have extracted the hour variable from the time factor

Hour :	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Instances :	29	26	27	24	26	24	39	92	120	157	175	28	26	73	196	188	144	24	25

Creating a new factor with Ideal fuel economy

As the ideal fuel_economy of the car

- City Mileage: 15-23 Kmpl
- Highway Mileage: 17-25 Kmpl
- Trimming on both extremes to get the most probable range 17 - 23

****** Placing myself in customer Shoe,I expect certain close range of values to get the best economy rather than sticking to one value ,As it is difficult to the driver to drive on single value for best performance.(The tachometer will also be in term of thousands,so it will be difficult to drive at single value).
That is the reason why I have used Ranges for easy and clean visualisation,so that we can draw underlying conclusions from the data. ******

The exact value provided by the company is 17.4

So I have created a new variable based on fuel economy of the instances where fuel_economy is greater than 17 as one level and less than 17 as another level.

Levels of economy:

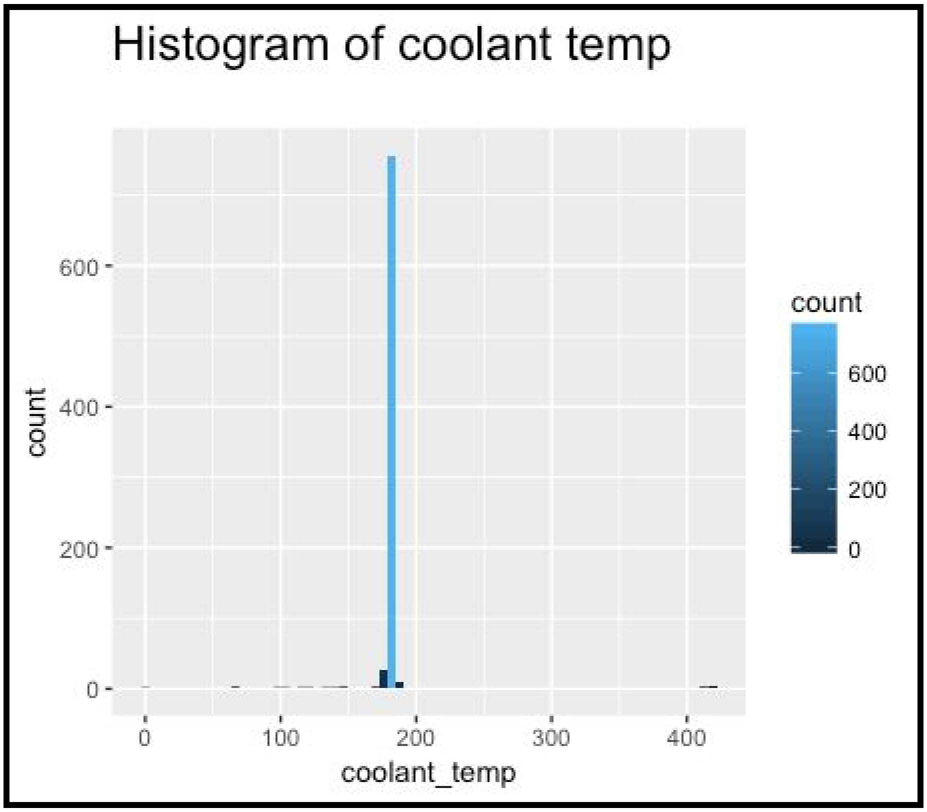
- **"Greater than 17"**
- **"less than 17"**

Values collected from various resources in the internet:

- Ideal speed: 50-60
- IDLE RPM:800-900 RPM
- Fuel consumption=17.4

Chapter 3:

Exploratory Data Analysis:



Explanation:

- Histogram of coolant temperature

Conclusion:

- As most of the data has only one value it does not influence the other variables.so I am not using this in further explanation.
- Most cars, the normal operating engine temperature is in a range of 195 to 220 degrees Fahrenheit. So its is working perfectly for this car.

IMPORTANT:

To understand the flow of report :

In Visualisation with respect to RPM(Method-1).

1. RPM vs Fuel_economy. (Uni-variate)
2. Speed for different RPM ranges.(Bi-variate)
3. RPM vs Speed vs Fuel_economy.(Multi-variate)

We have covered single factor contributing fuel_economy and combination of factors contributing fuel_economy and have drawn conclusions.

Dividing the data in two parts (economy >17,economy<17) and performing with respect to RPM of Divided data.

Part-1-(economy >17)

4. RPM for which fuel_economy >17.
5. Speed for the above RPM values (RPM for which fuel_economy >17)

The same is repeated for RPM with economy <17

Part-2- (economy<17)

6. RPM for which fuel_economy <17.
7. Speed for the above RPM values (RPM for which fuel_economy <17)

After explaining both parts.

8.I have explained a graph of two regions(economy>17 and Economy<17) in single graph to summarize what we have concluded from the above graphs.

Summing up all the factors contributing to both the parts(economy>17 and Economy<17) individually(and drawing conclusions

***** The same steps are followed with Visualisation with Speed (Method-2)**

Difference between two methods:

In Method-1(In Visualisation with respect to RPM)-I have used factors contributing with respect to RPM and went deeper by filtering the RPM(RPM for fuel_economy>17 and RPM for fuel_economy <17 and their speeds association with each)**.**

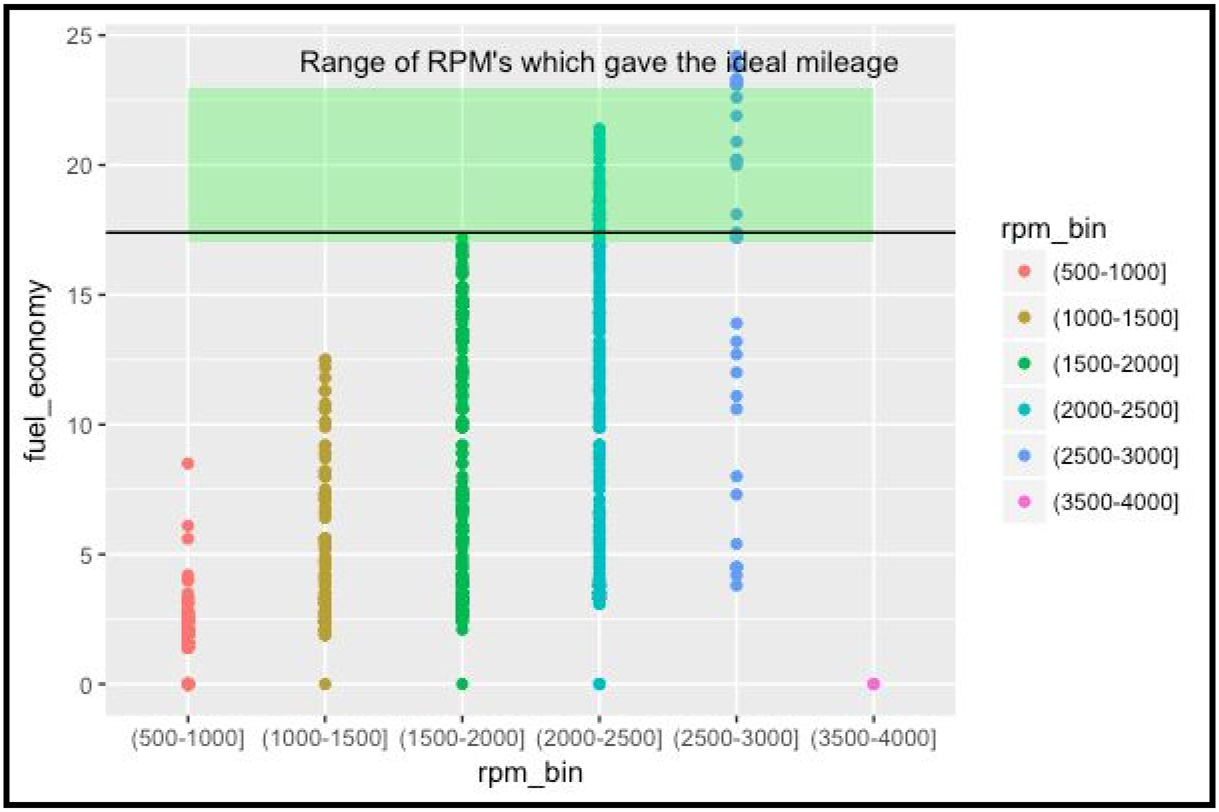
In Method-2(Visualisation with respect to speed)-I have explored factors contributing with respect to speed and went deeper by filtering the Speed (Speed for fuel_economy>17 and Speed for fuel_economy <17 and their RPM's association with each)

Flowchart of methods followed:

Visualisation with respect to RPM	Visualisation with respect to Speed
Method-1	Method-2
RPM vs Fuel economy	Speed vs Fuel economy
↓	↓
Distribution of RPM in Total data	Distribution of Speed in Total data
↓	↓
RPM vs Speed	Speed vs RPM
↓	↓
RPM vs fuel_economy vs Speed	Speed vs fuel_economy vs RPM
↓	↓
(Part-1)	(Part-1)
RPM of data(Fuel_economy >17)	Speed of data(Fuel_economy >17)
↓	↓
Speed for the above RPM range	RPM for the above Speed range
↓	↓
(Part-1)	(Part-2)
RPM of data(Fuel_Economy <17)	Speed of data(Fuel_Economy <17)
↓	↓
Speed for the above RPM range	RPM for the above Speed range
↓	↓
Two different parts in single graph	Two different parts in single graph

Method -1:

Visualising with respect to rpm:



Explanation:

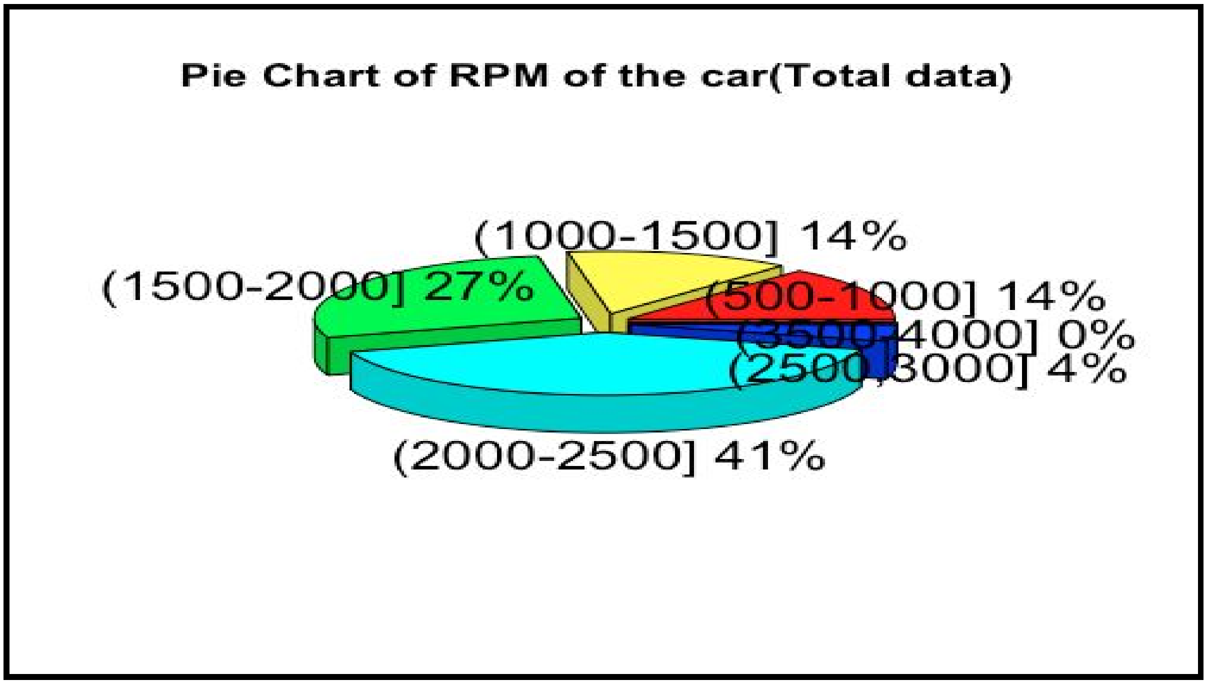
- The above green shaded explains the RPM ranges which has given the ideal mileage range of **17<= economy < 23**
- The horizontal line indicates the exact mileage given by the Honda company **[17.4]**

Conclusions:

From the above graph we can conclude that

- RPM ranging between 2000-3000 has given the fuel_economy between 17 and 23.
- RPM range 1500-2000 has given only once in the collected data.So I am considering only the highly probable cases.

Pie chart explaining the RPMs in which car was driven



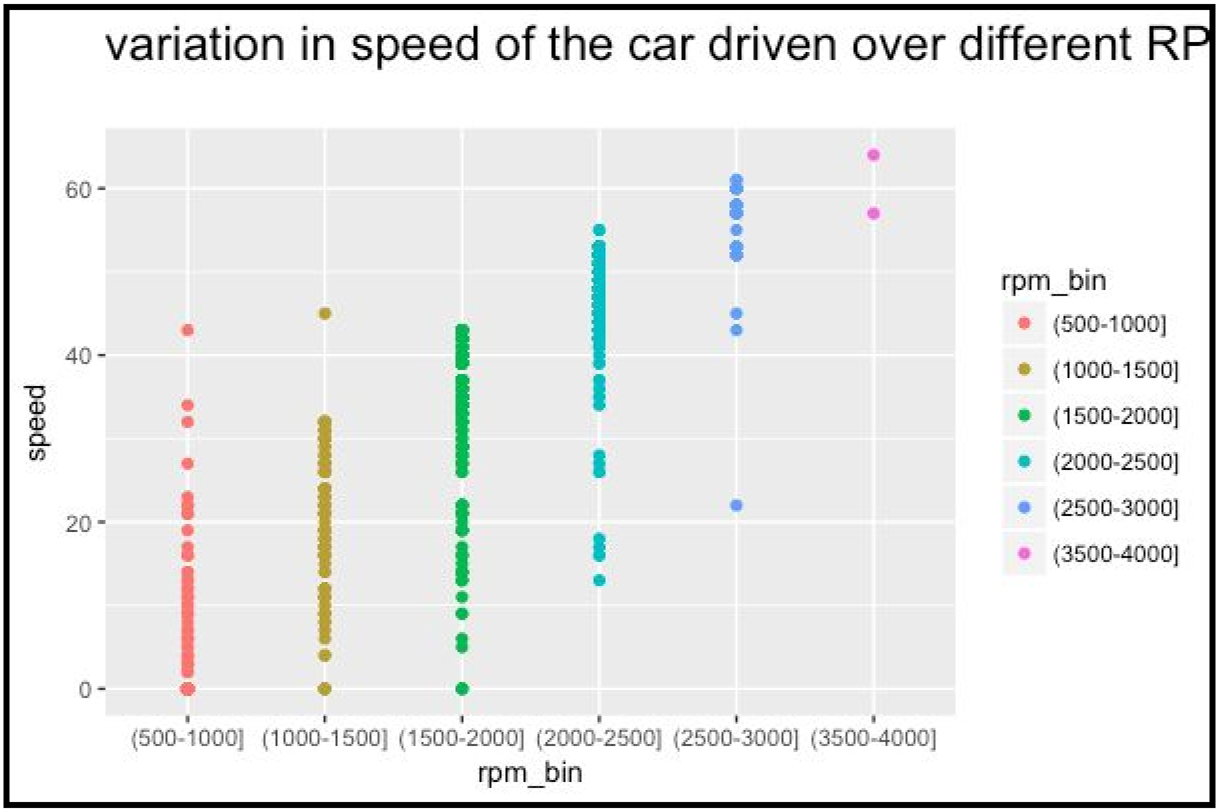
Explanation:

- The above graph explains ,over the collected data in what RPM ranges the car was driven and how often the car was driven in these RPM ranges

Conclusions:

From the above graph we can conclude that

- Nearly 50 % of the time the car has been driven in range of 2000-2500
- 82 % of the time the car has been driven with in 1000-2500 RPM ranges

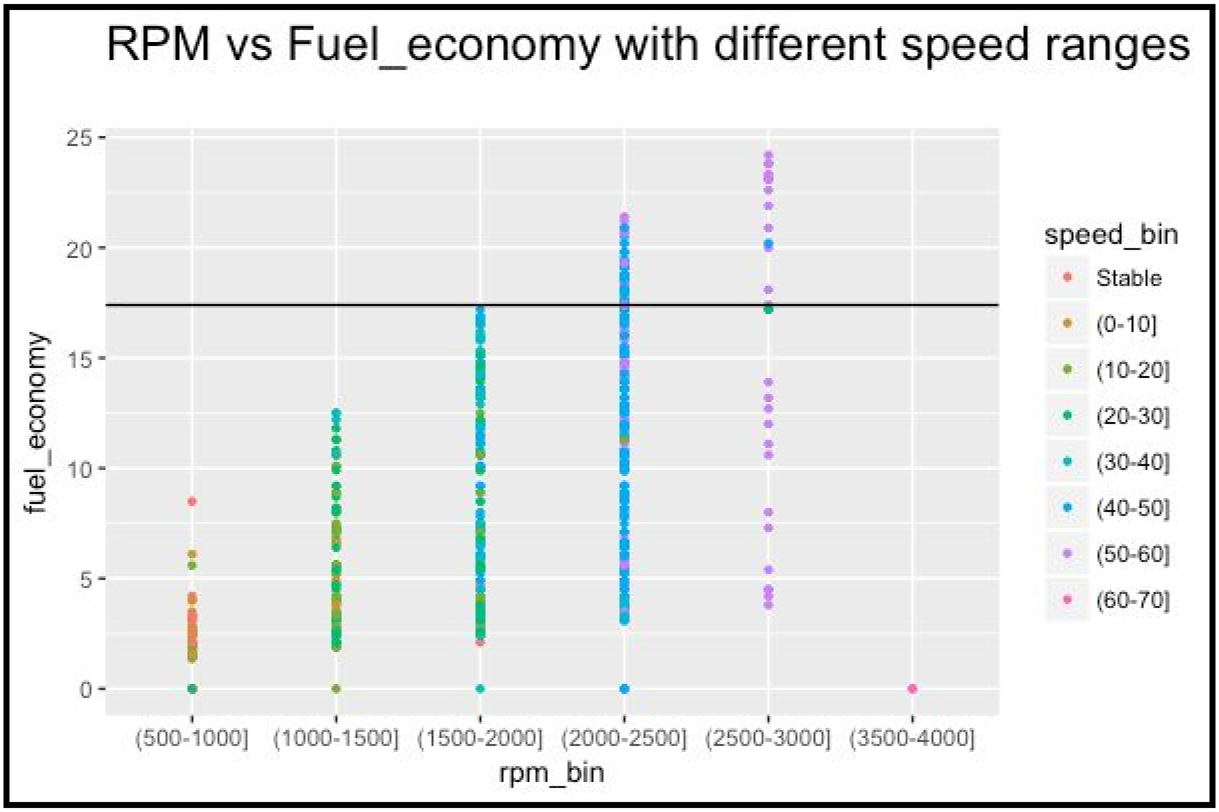


Explanation:

- It explains the speed attained by the vehicle in different RPM ranges.

Conclusion:

- We can notice that with in rpm range of (500-1000] the car has attained the speed of 42,as long you're not stalling the car from the low engine speed there is no harm. some engines don't work very well below 1500 rpm and may stall, which, if done repeatedly, may have a negative impact on your engine.



Explanation:

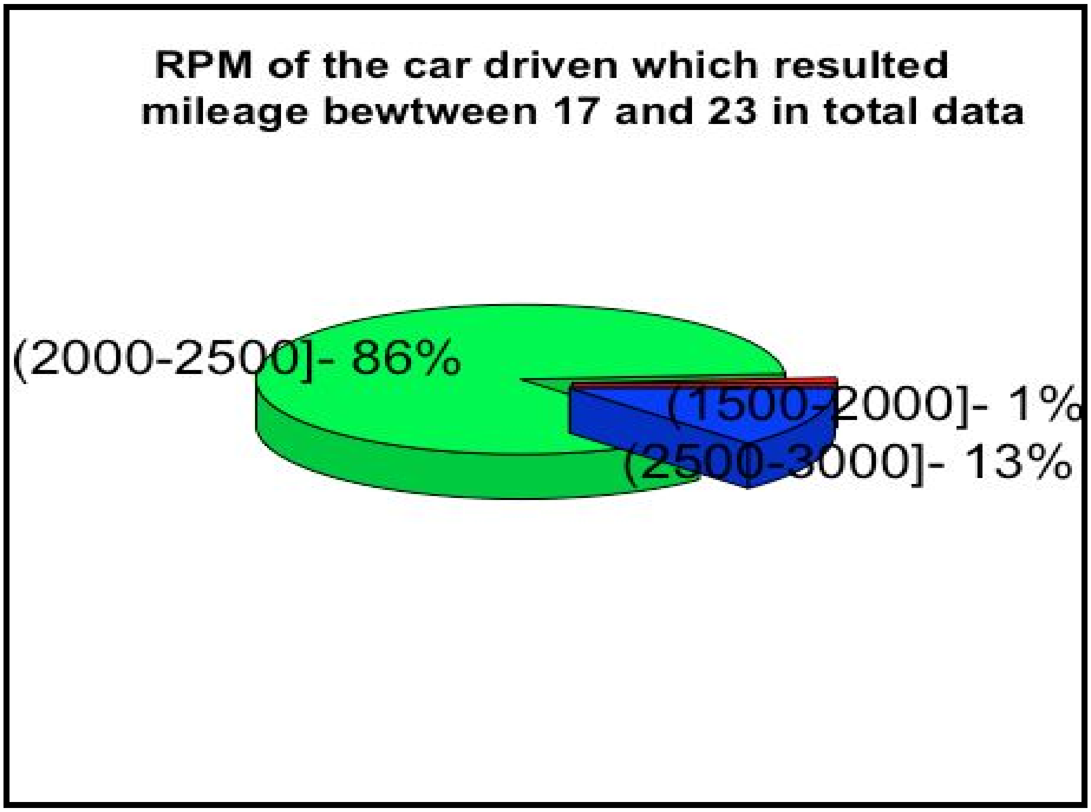
- This graph explains the fuel_economy attained by the car from different RPM's and speed ranges
- The line explain the cut off point of 17.4 as mentioned by the company.

Conclusions:

- We can notice that only speed ranging from 40-60 has given the mileage greater than 17.4 with rpm range of 2000-3000

As this graph is difficult to interpret I have made following graphs explaining in detail.

Part-1: Fuel_economy >=17



Explanation:

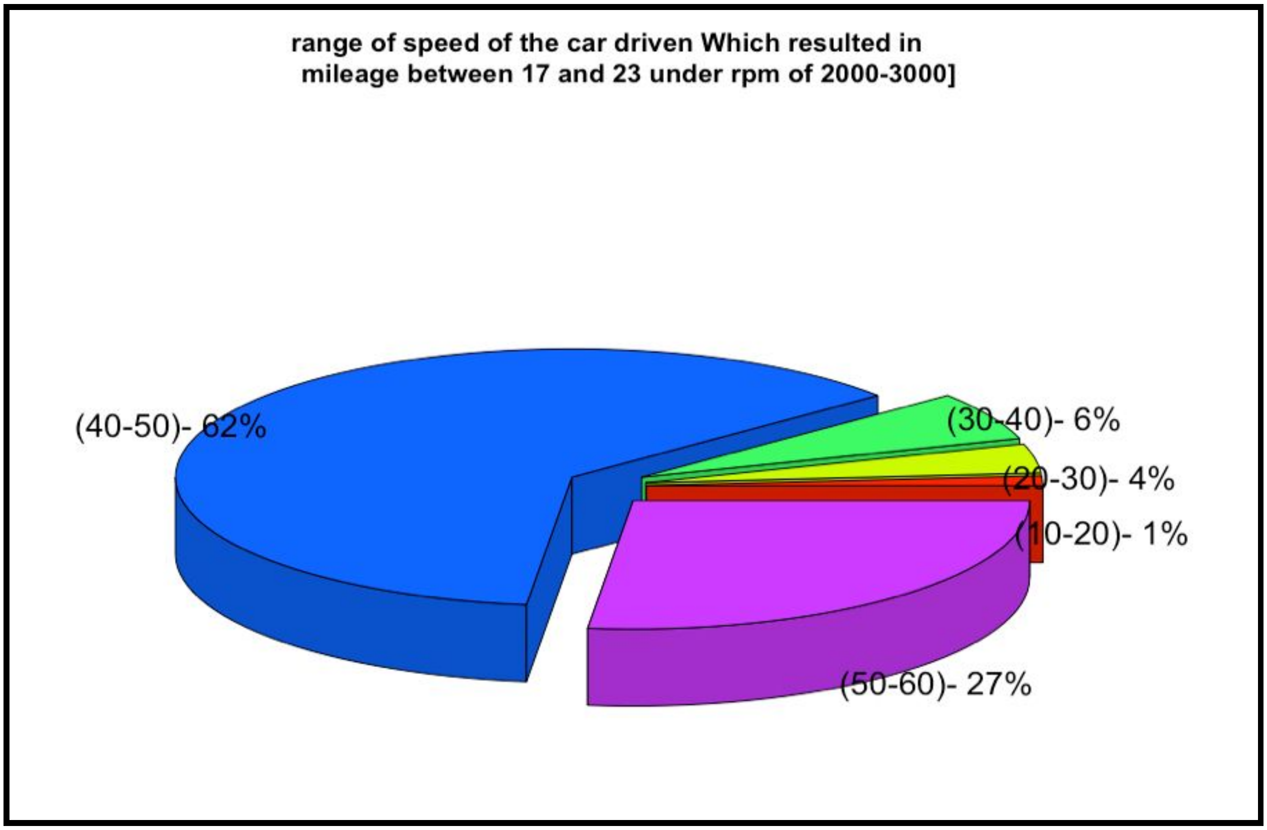
- This graph explains the how often the RPM ranges ,with which the car has attained the mileage between 17 and 23 .

Conclusions:

- We can conclude that over the instances collected with fuel_economy >17 ,86% of the time the RPM meter is between (2000-2500].
- The car has been absolutely performing well in the rpm range of (2000-3000].

As we have seen 86% of the time the car with RPM 2000-2500 has given the mileage between 17 and 23,

So let's observe what was speed attained by the car in RPM range 2000-2500 with mileage between 17 and 23 ?



Explanations:

- This graph explains speed used by the car in RPM range of 2000-3000 to obtain fuel_economy between 17 and 23

Conclusions:

- 89 % of the time the car was in the speed range of 40-60 with RPM range of 2000-3000 which resulted in mileage between 17 and 23. This shows what contributed for mileage between 17 and 23.

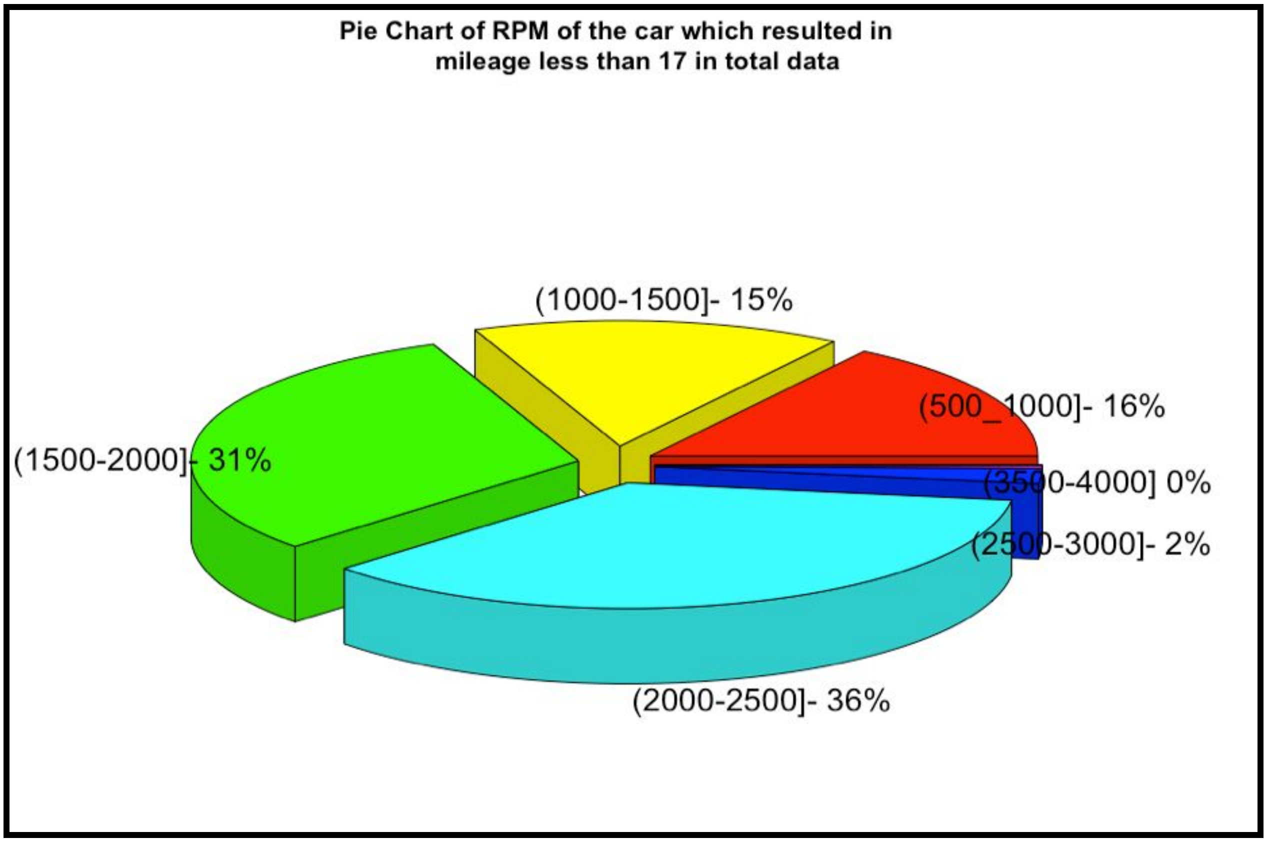
Important conclusion drawn from the above graph to analyse the behaviour of the car:

Car has given mileage between 17 and 23 with

- RPM:2000-3000
- Speed:40-60

As we have analysed the optimum performance of the car, let's analyse what is the reason for car giving mileage less than 17.

Part-1: Fuel_economy<17



Explanation:

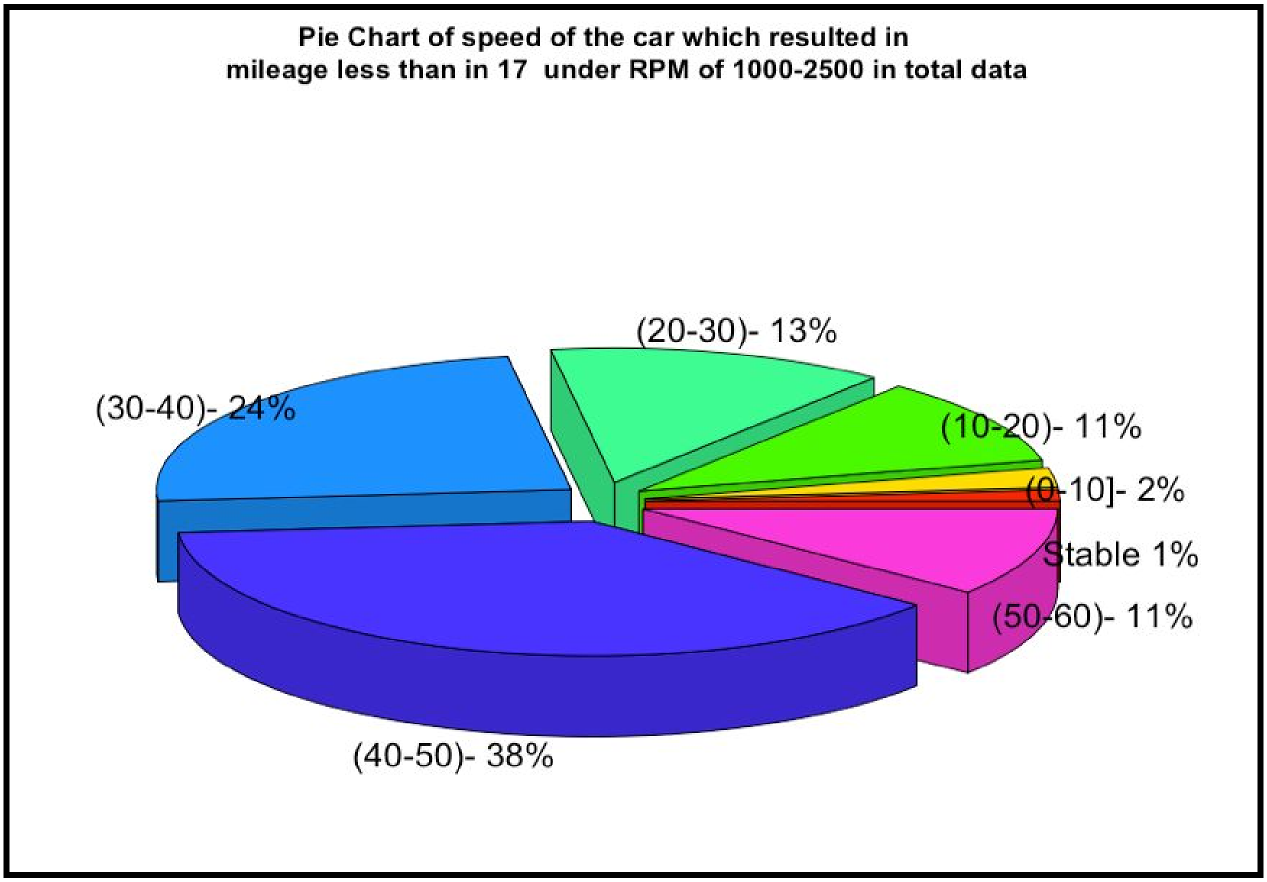
- This graph explains RPM ranges for which the fuel_economy is less than 17

Conclusion:

- We can notice that RPM range of 1000-2500 has been contributing more (83%) for decrease in fuel_economy.

As we have seen that RPM 1000-2500 is contributing for fuel_economy less than 17,

Let's see what was the speed attained by the car in 1000-2500 RPM range which resulted in fuel mileage less than 17 ?



Explanation:

Speed attained by the car with RPM between 1000-2500 which resulted in fuel_economy less than 17

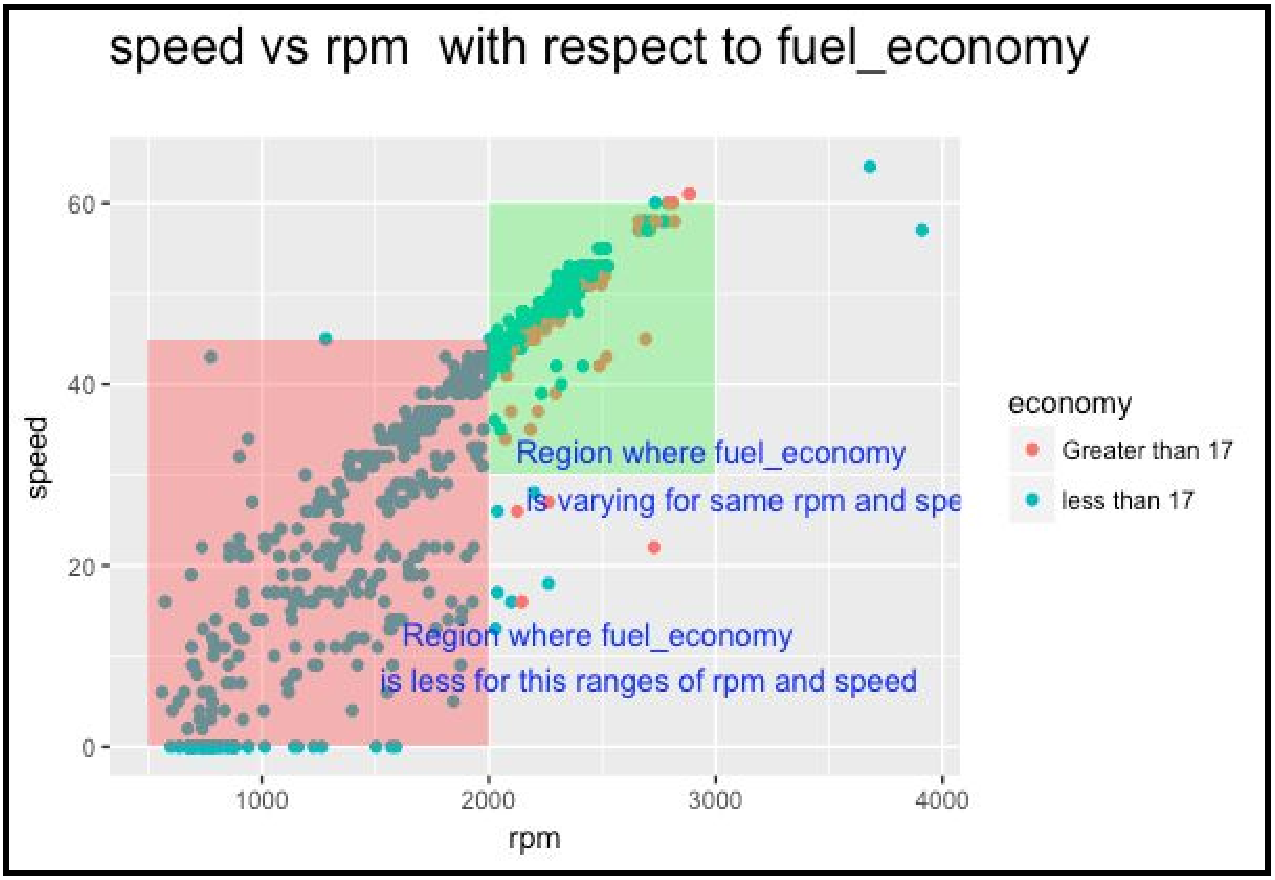
Conclusions:

We can notice that speed range of 30-50 has contributing more for decrease in mileage

(<17) with RPM 1000 and 2500.

Clashes between Mileage <17 and mileage <17 for same RPM and Speed ranges.

Let's see the region where there is a clash between fuel_economy>=17 and fuel_economy<17 graphically



Explanation:

The above plot explains the regions where the car has given the ideal mileage >= 17 and less mileage <17 .

Conclusions:

Red region:

The red plot has only blue dots(instances for which the fuel_economy is < 17.

So if we ride the car in this region it will give fuel_economy <17

For fuel_economy <17

- speed=0 to 45
- RPM=0 to 2000.

We can conclude car was driven in above mentioned ranges which resulted in fuel_economy less ideal (<17).

Green region:

The green region explains when the car has given ideal mileage >=17 and well as less ideal mileage <17

There is a possibility of two chances if we ridge in green region.

Why is that happening?

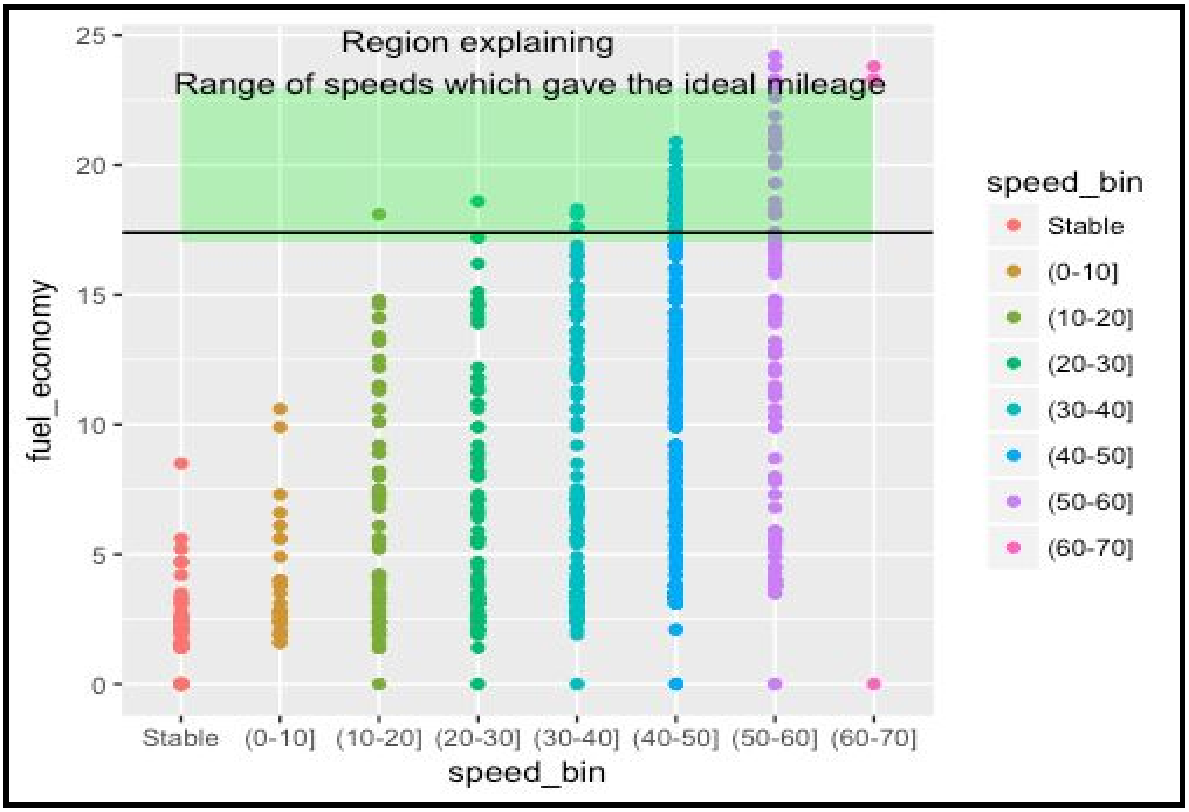
RPM-stands for Revolutions Per Minute or the speed of the engine. The faster it revolves the faster and more numerous the combustion cycles and hence the fuel burnt.The lower the gear, the higher the rpm. The higher the rpm, the more torque the engine is producing, and the more fuel it is using

Possibility:

- Driving at a high speed in a lower gear
- Driver in driving aggressively.
- Badly worn spark plugs and clogged or dirty filters can decrease engine performance, which will result in decreased gas mileage and performance.

Method-2:

Visualisation with respect to Speed:



Explanation:

Graph explaining which speed ranges has given the fuel_economy>=17.

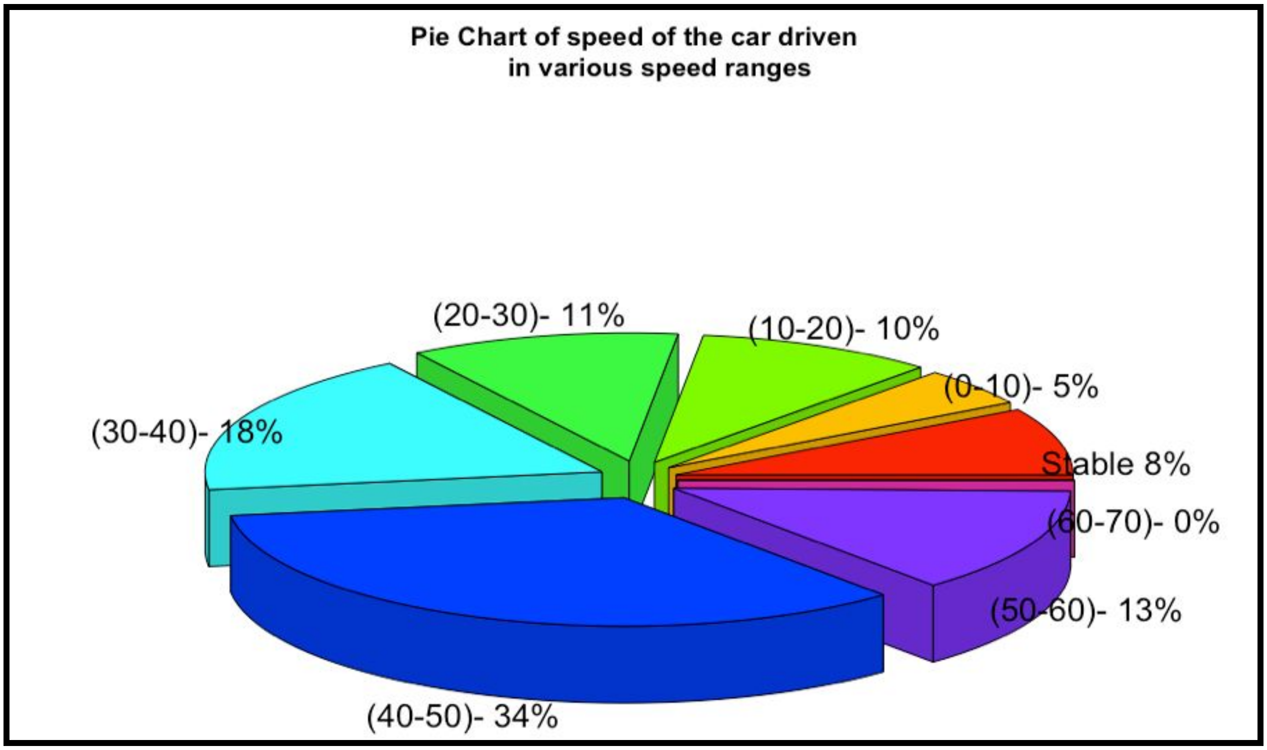
Line is the cut off line of fuel_ecomomy=17.4 mentioned by the company.

Conclusion:

We can notice that speed ranging from 20-70 has given the fuel_economy <17 and less than 23.

The speed ranging from 50 to 70 has given more than the ideal fuel_economy (fuel_economy>23) which is an important point to consider.

Let's see how often the car was driven in these speed ranges graphically



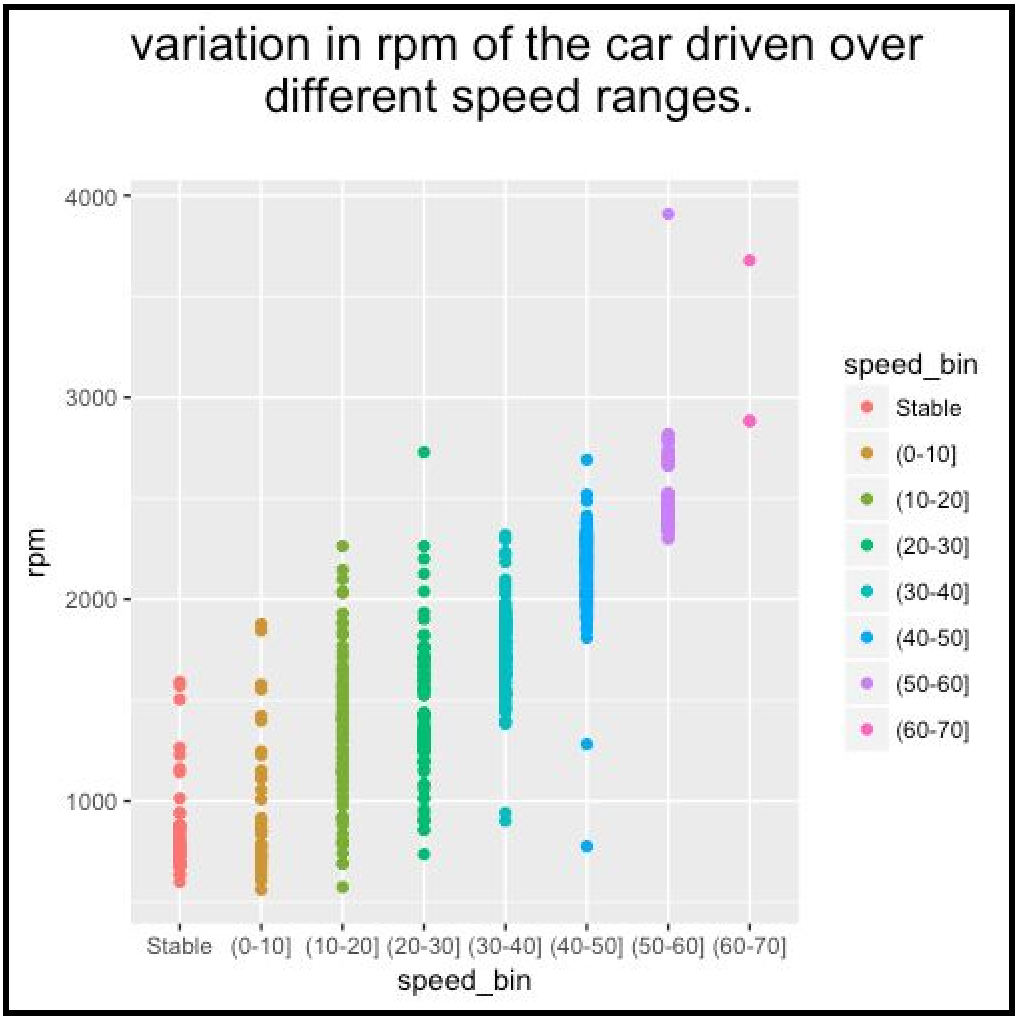
Explanation:

- The above graph explains how often the car has been driven in different speed ranges

Conclusion:

- Most of the time the car has driven between the speed 30-60
- Over the data collected the speed of the car was between 60-70 is less than 1%

Lets see what was the range of the rpm in different speeds?

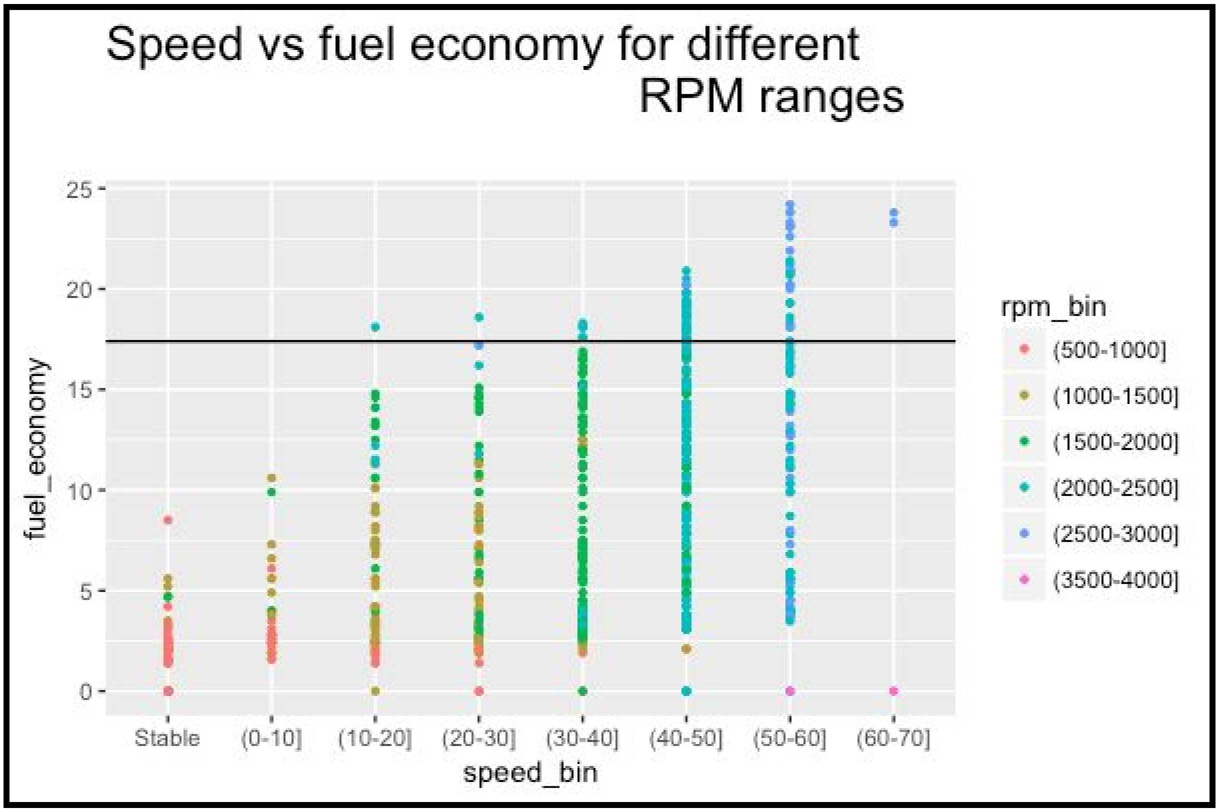


Explanation:

This graph explains the RPM of the vehicle in various speed ranges.

Conclusion:

- The highest RPM is attained in speed range of 50-60.
- The lower the gear, the higher the rpm. The higher the rpm, the more torque the engine is producing, and the more fuel it is using. Automatic transmissions take some of this control out of the driver's hands, but they, too, can be manipulated to maximize fuel efficiency



Explanation:

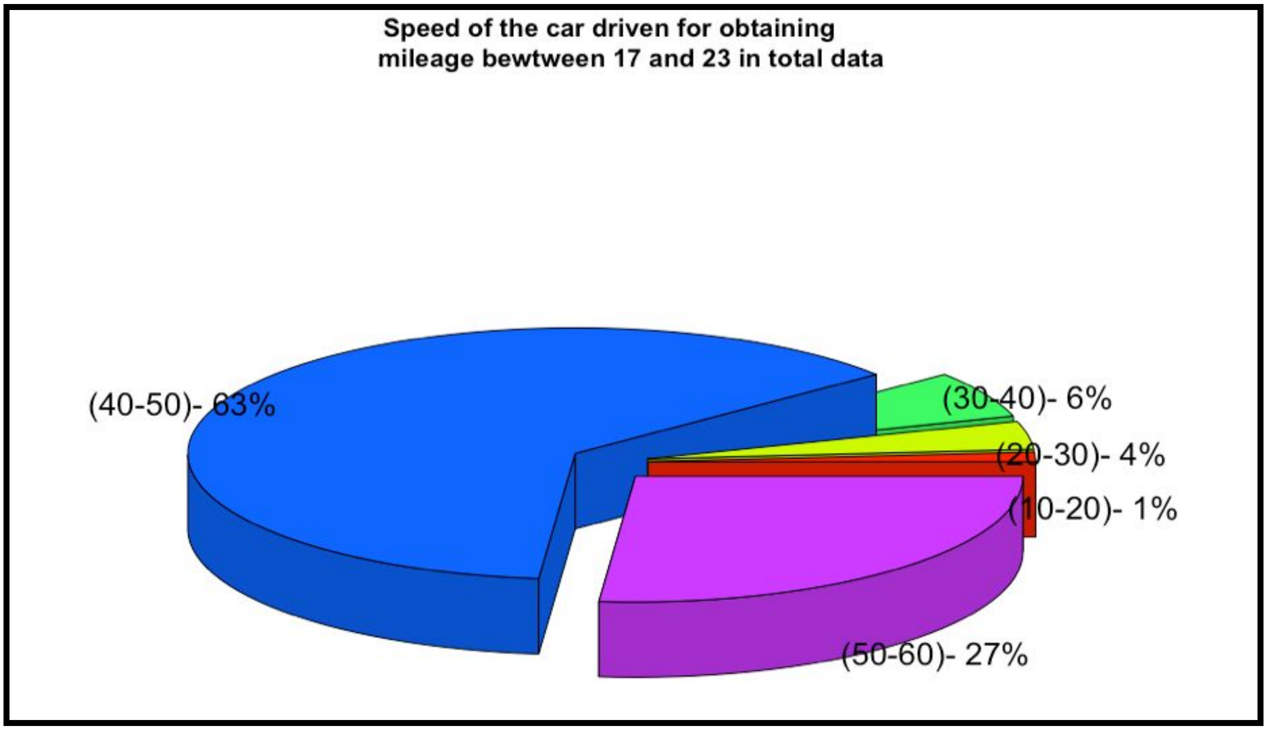
- This graph explains the range of fuel economy for different speed and RPM ranges.

Conclusion:

- From the graph we can notice that RPM ranging between 2000 and 3000 has only given the fuel_economy >17

This graph is very difficult to interpret so let's plot more visually to explain the RPM for fuel_economy >=17

Method 1-Part-1:(Economy >17)



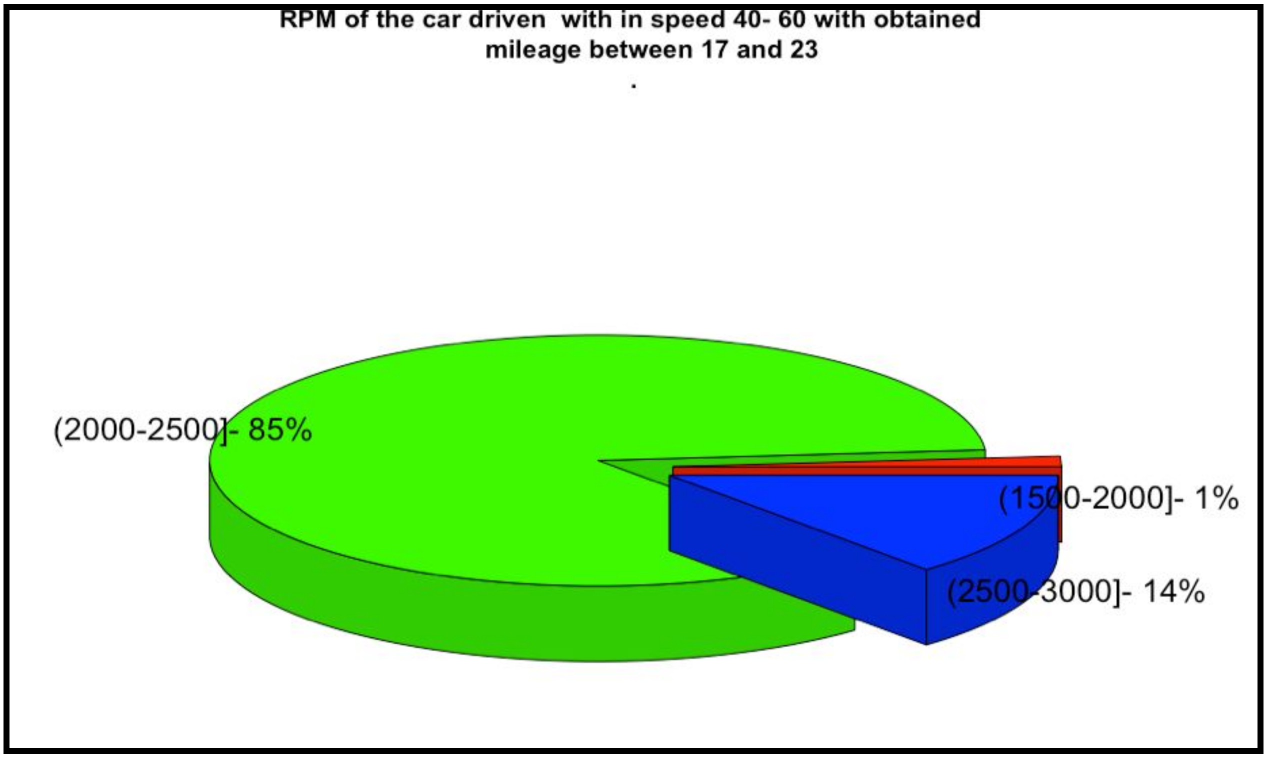
Explanation:

- over the data collected for what range of speeds that car has given the mileage between 17 and 23

Conclusion:

- We can notice that almost 90% of the time the car has been driven in range of 40-50 which resulted in mileage > 17

Let's observe what was the RPM of the vehicle in speed range of 40-60



Explanation:

- This graph explains the rpm used by the vehicle with speed range of 40-60 which resulted in mileage between 17 and 23.

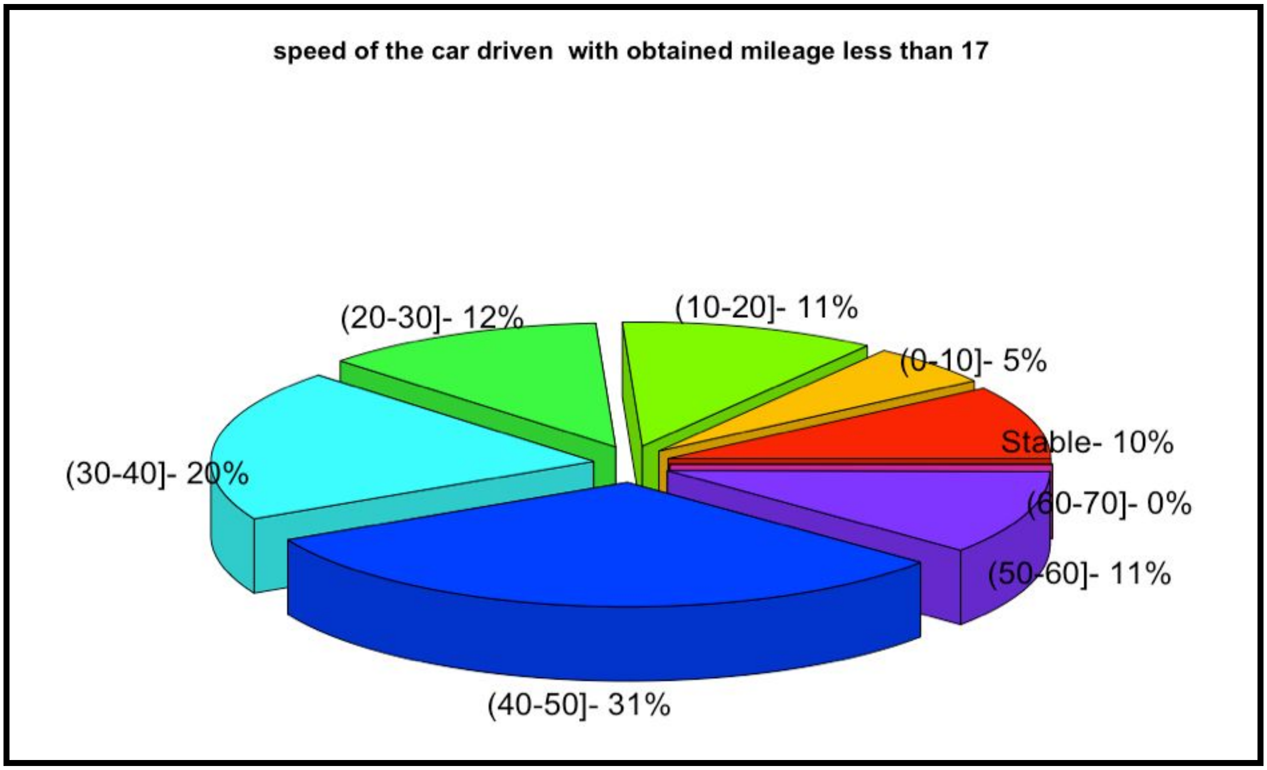
Conclusion:

- Almost 100% of the time the car was in range of 2000-3000 RPM with speed 40-60 which resulted in mileage >17 and <23.

As we have observed what are the RPM and speed ranges which resulted in mileage between 17 and 23.

Let's observe what range of values have resulted in mileage less than 17.(<17)

Method 2-Part-2-(Economy <17)



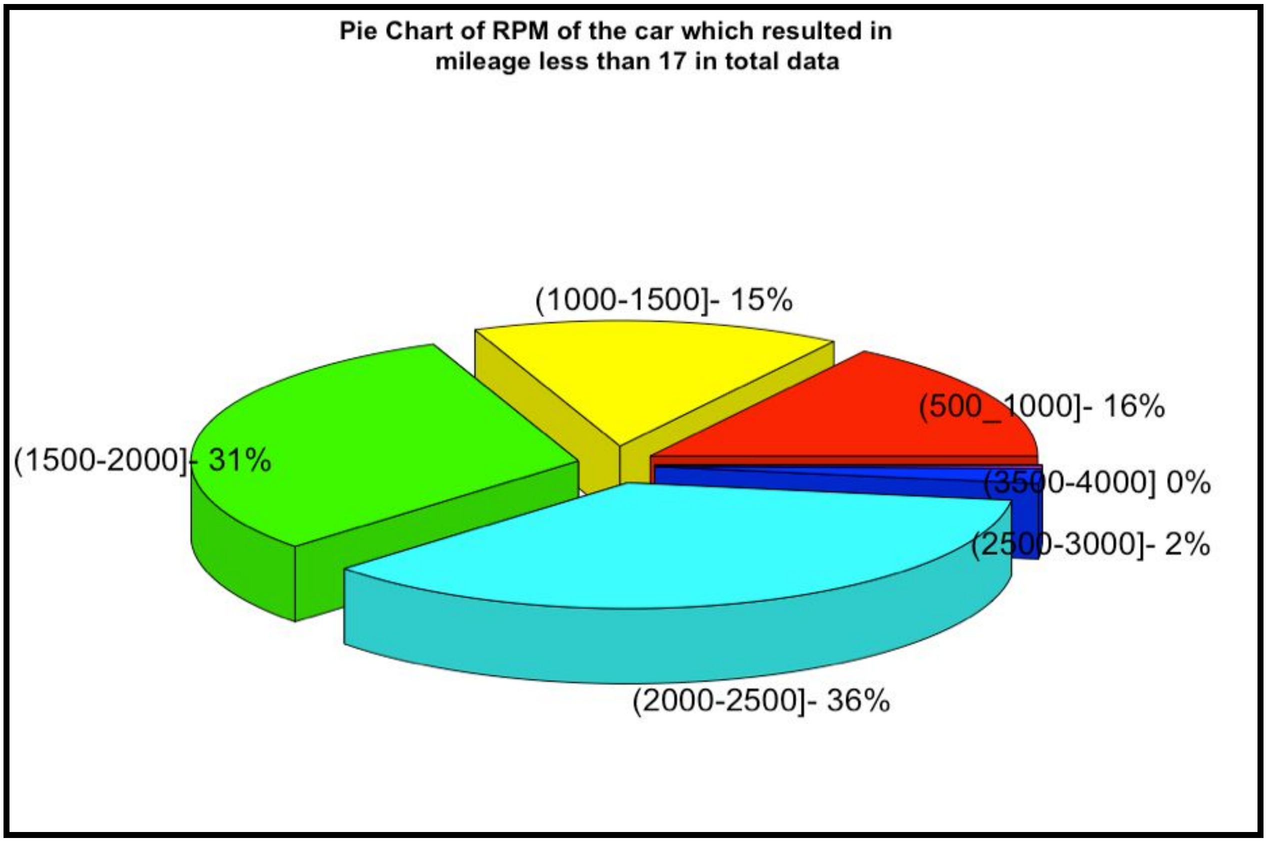
Explanation:

- This graph explains the different speeds attained by the vehicle which resulted in fuel_economy less than 17.

Conclusion:

- We can conclude that the all range of speeds has been a part which resulted in mileage less than 17 .

So Let's look at the RPM values for all ranges of speed which resulted in fuel_economy <17



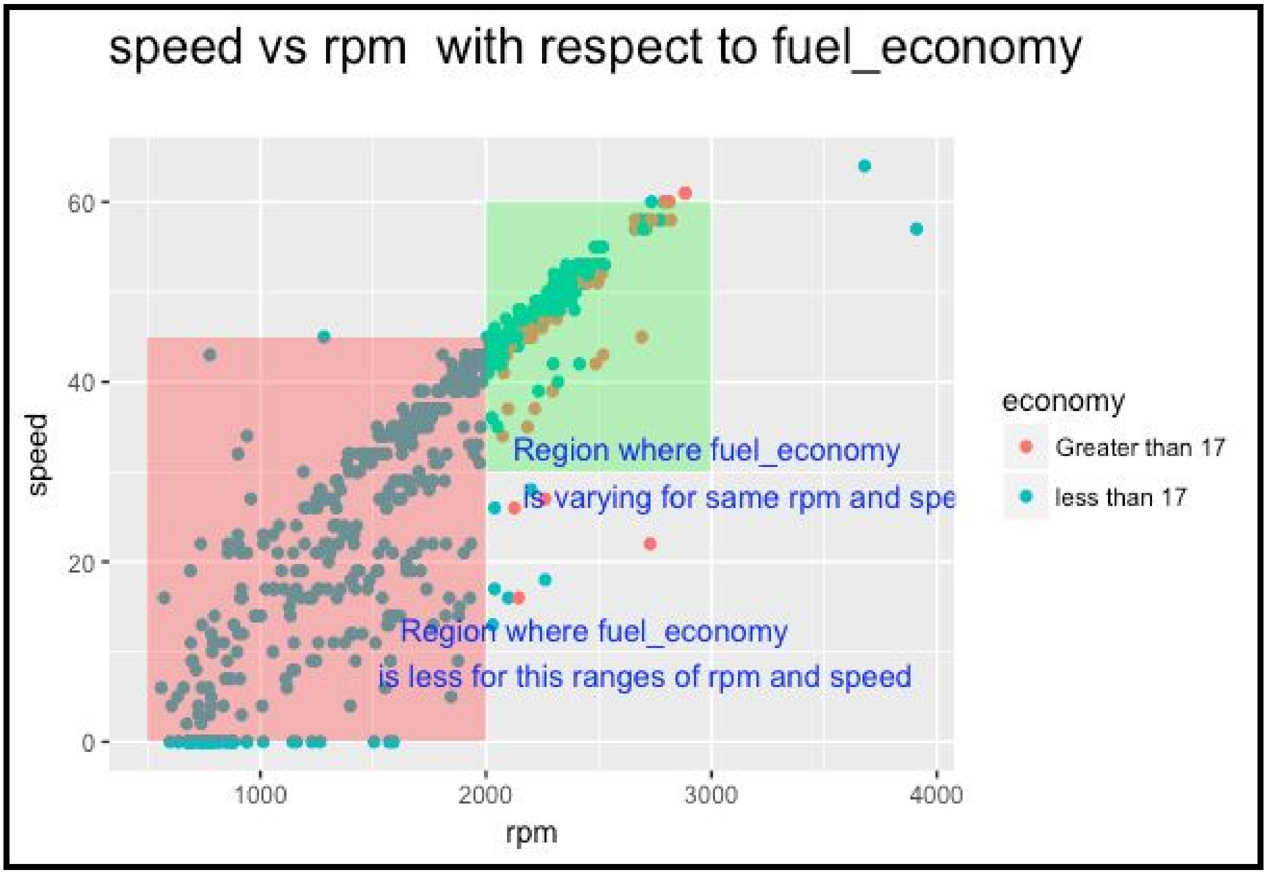
Explanation:

RPM of the vehicle which resulted in decrease of mileage(<17) across all speed

Conclusion:

We can conclude that >75% of the time the RPM range of 1000-2500 has contributed for decrease in mileage across all speeds.

Let's observe the regions of performance differentiating the fuel_economy for different range of RPM and Speed values:



Explanation:

The above plot explains the regions where the car has given the ideal mileage >= 17 and less mileage <17 .

Conclusions:

Red region:

- The red plot has only blue dots(instances for which the fuel_economy is < 17.
- So if we ride the car in this region it will give fuel_economy <17

For fuel_economy <17

- speed=0 to 45
- RPM=0 to 2000.

We can conclude that is we ride the car in the above mentioned ranges the fuel_economy will be 100% less ideal (<17).

Green region:

- The green region explains when the car has given ideal mileage >=17 and well as less ideal mileage <17.
- There is a possibility of two chances fuel_economy greater and less than 17,if we ridge in green region.

Why is that happening?

RPM-stands for Revolutions Per Minute or the speed of the engine. The faster it revolves the faster and more numerous the combustion cycles and hence the fuel burnt**.**The lower the gear, the higher the rpm. The higher the rpm, the more torque the engine is producing, and the more fuel it is using

Possibility:

- Driving at a high speed in a lower gear
- Driver in driving aggressively.
- Badly worn spark plugs and clogged or dirty filters can decrease engine performance, which will result in decreased gas mileage and performance.

Some Key observations from all the graphs explained above:

General observation:

- 82 % of the time the car was driven with in 1000-2500 RPM ranges
- Most of the time the car was driven between the speed 30-60
- The instances where speed of the car between 60-70 is less than 1%(very less)

Factors which has fuel_economy >17

- RPM ranging between 2000 and 3000 has only given the fuel_economy >17
- speed ranging from 40-60 has given the mileage greater than 17.4 with rpm range of 2000-3000.

Factors which resulted in fuel_economy <17

- To be precise speed of 0-45 and rpm of 0-2000 combination has given the speed less than 17.

- Car driven with Speed range of 30-50 resulted in for decrease in mileage (<17) under RPM 1000 and 2500.

Conclusions from the above factor for fuel_economy > 17.

1. RPM-range-2000-3000
2. Speed Range- 40-60
3. Be as smooth as possible with the car's controls. Treat the accelerator and the brake pedals with respect.
4. Rough speedy acceleration and sudden brakes might be more fun on the road but harms your vehicle's fuel efficiency.
5. A car consumes more fuel in lower gears when accelerating. If the engine's not pulling, shift down instead of pressing the throttle
6. Don't drive your car with dirty or clogged filters, such as the air filter or the oil filter if you don't want to shell out at the pump. Get them cleaned or replace them.

Conclusions from the above factor for fuel_economy < 17.

1. RPM range 1000-2500.
2. speed=30-50.
3. Driving aggressively.
4. Having bad oxygen sensors and air filters can reduce your gas mileage by up to 20%.
5. spark plugs are responsible for sparking combustion in your engine. If they misfire, or are working poorly, this can affect your gas mileage in a negative way.

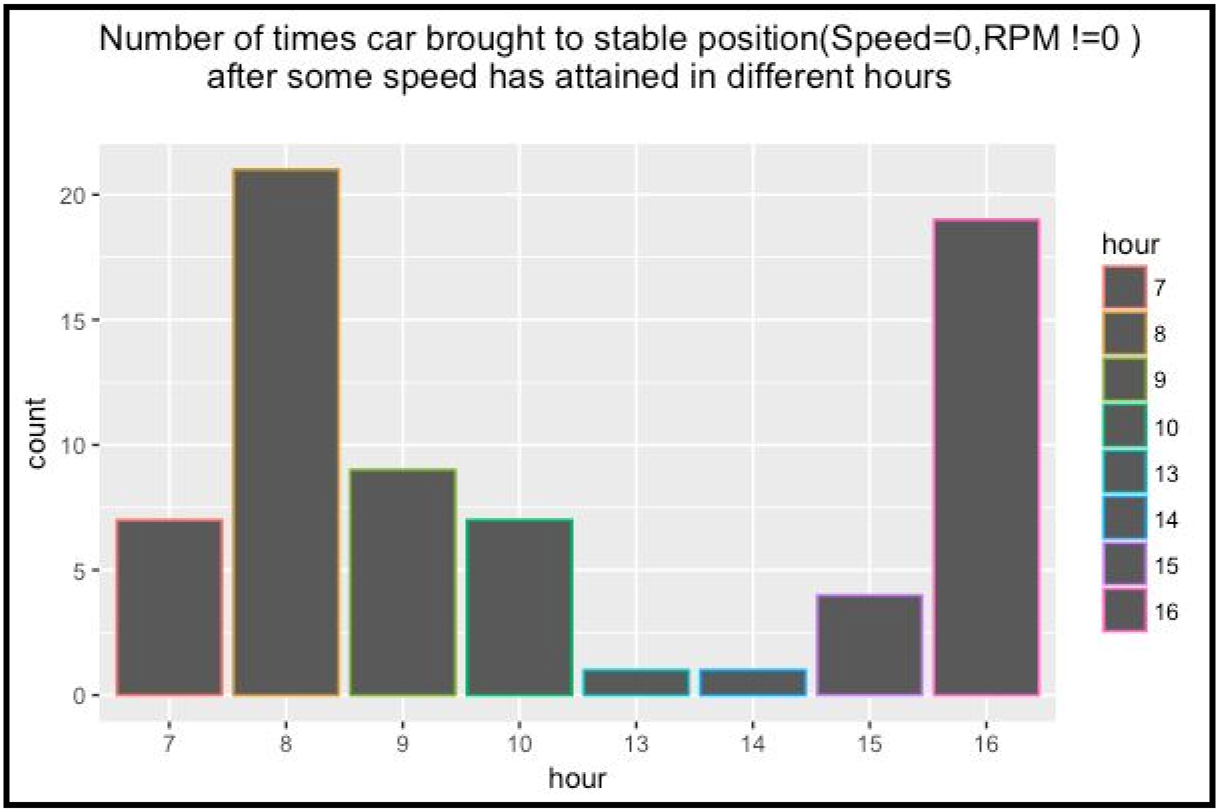
**** If we look at the above two conclusion there are some cases for which the car has performed good and bad for same range of values. ****

Reason?

This can lead to many possibilities.

- Aggressive driving
- Improper functioning of throttle
- Gear shifting
- Pressing throttle aggressively

Condition when Speed=0 but RPM =! 0



Explanation:

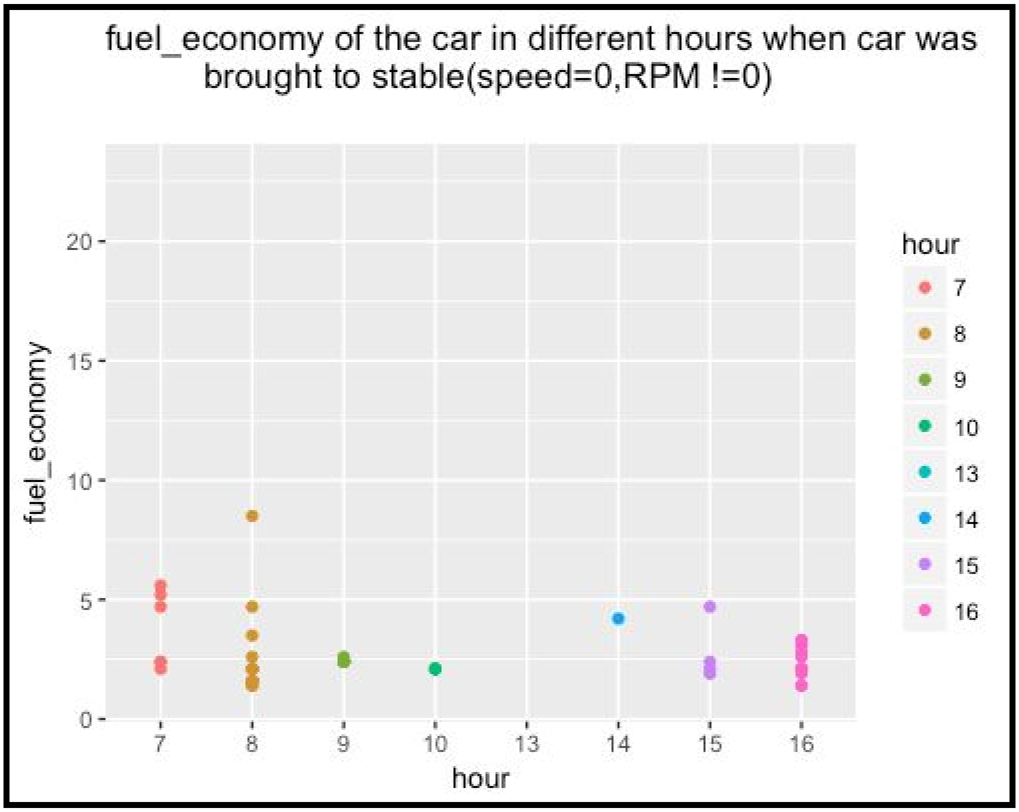
- Graph explaining how many time the car has brought to neutral state or stable state once it has attained some speed(after starting)

Conclusion:

- The lower the gear the higher the RPM and more the fuel is burnt.If we stop in between we should start moving in lower gear which is of high RPM and this decreases the fuel efficiency of the engine.

Ex: Fuel_economy is traffic will be very less

Let's observe

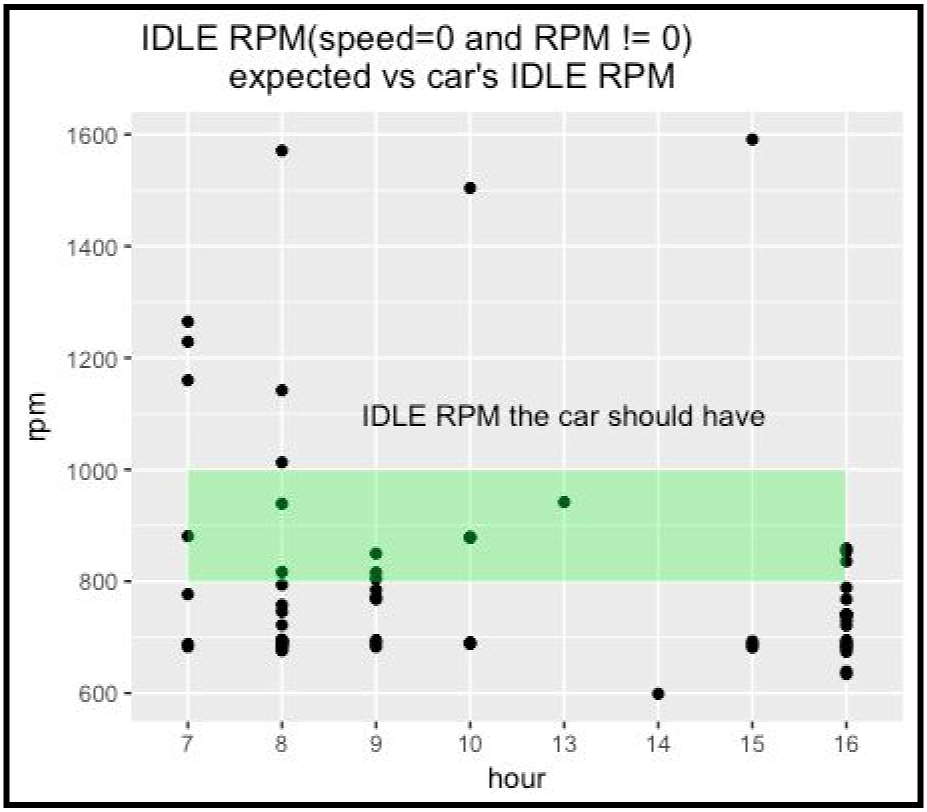


Explanation:

- This graph explains the economy of the instances when the car is being brought to stable.

Conclusion:

- From the above graph we can say that if we frequently being the car to stable it will lead to reduction in mileage as you have to gain speed again with low gear which has high RPM, so more burning of fuel.



Explanation:

- IDLE RPM of the car in different hours

Conclusion:

- The green region explaining the IDLE RPM the car should have
- The above graph also shows there has been instances where car's IDLE rpm exceed the general range of the IDLE RPM.

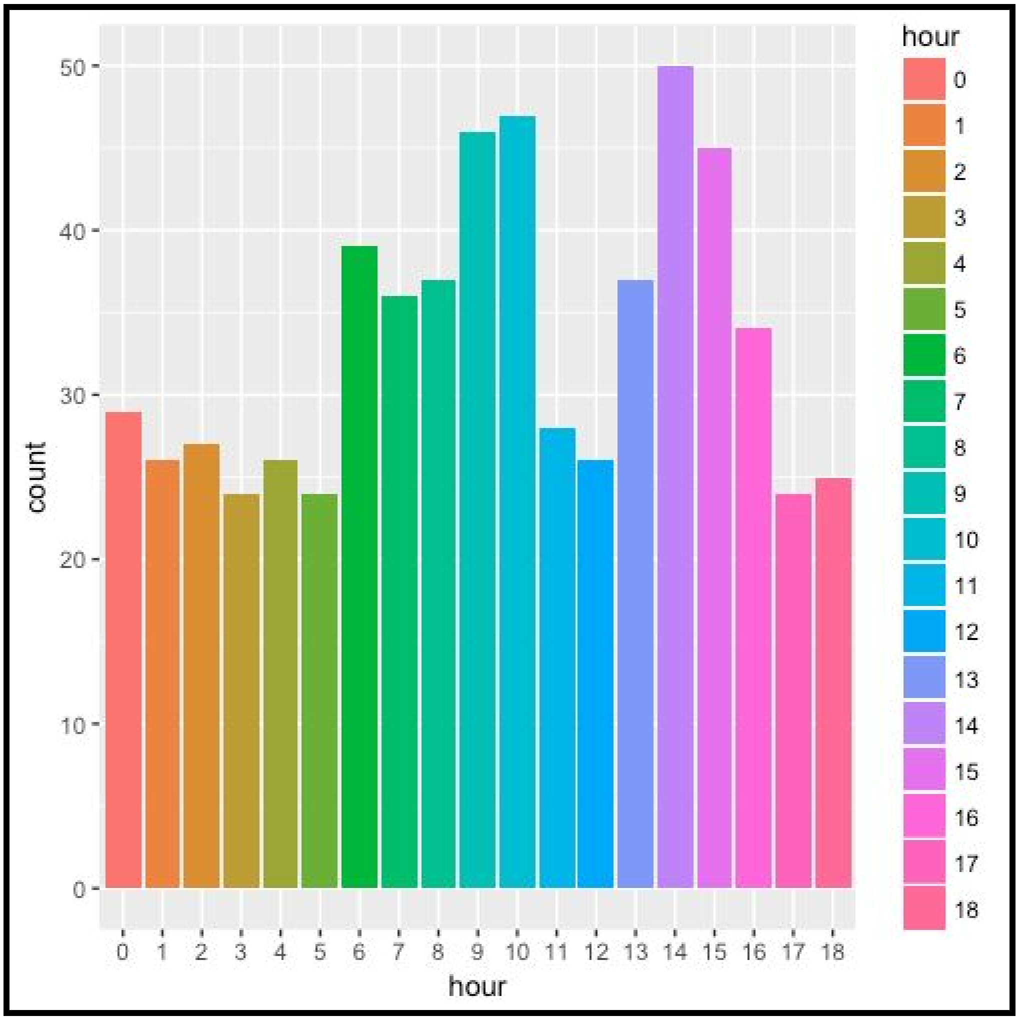
Reasons:

- **Vacuum Leak**
- **Faulty Idle Control Valve (ICV)**
- **Faulty Fuse**
- **Malfunctioning Throttle**

Condition when Speed=0 and RPM = 0:

Number of instances in each hour where speed=0 and RPM=0,

Key has inserted (power supply will be on[through battery]) but engine will be in off condition.



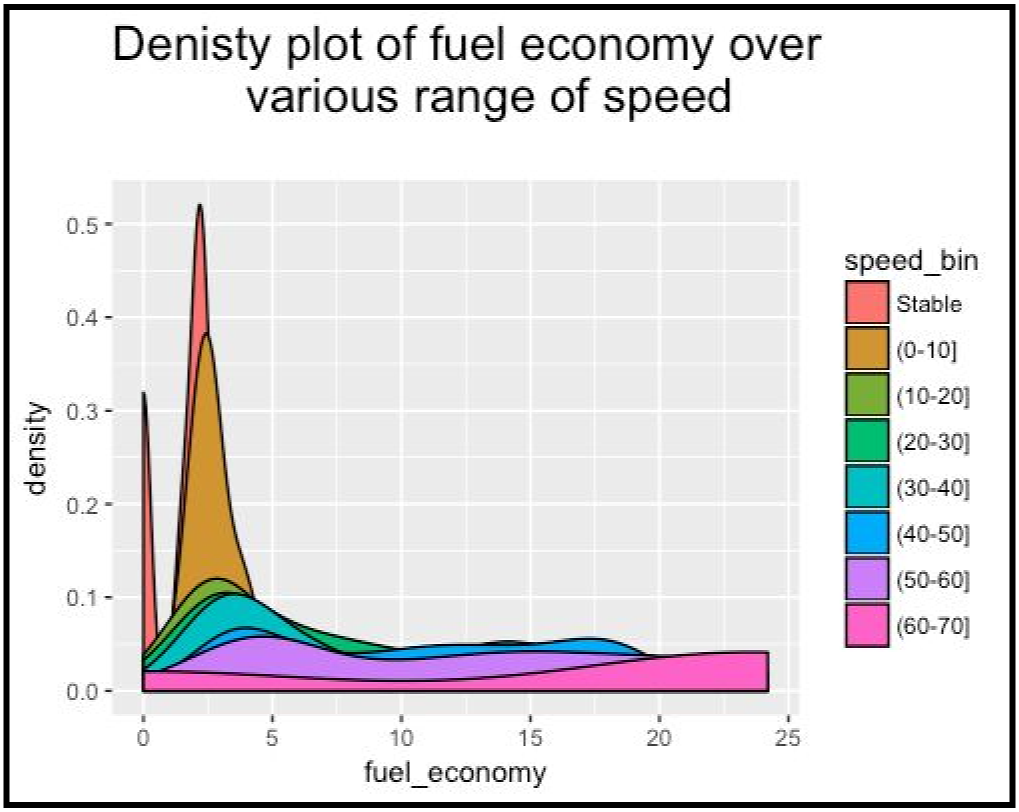
Explanation:

- This graph explains the count in each hour where speed=0=RPM

Conclusion:

- It wastes more fuel starting your car than you would by idling.
- This leads to bad performance of fuel_economy and Engine
- This leads to more draining of battery power.

Extra figures :

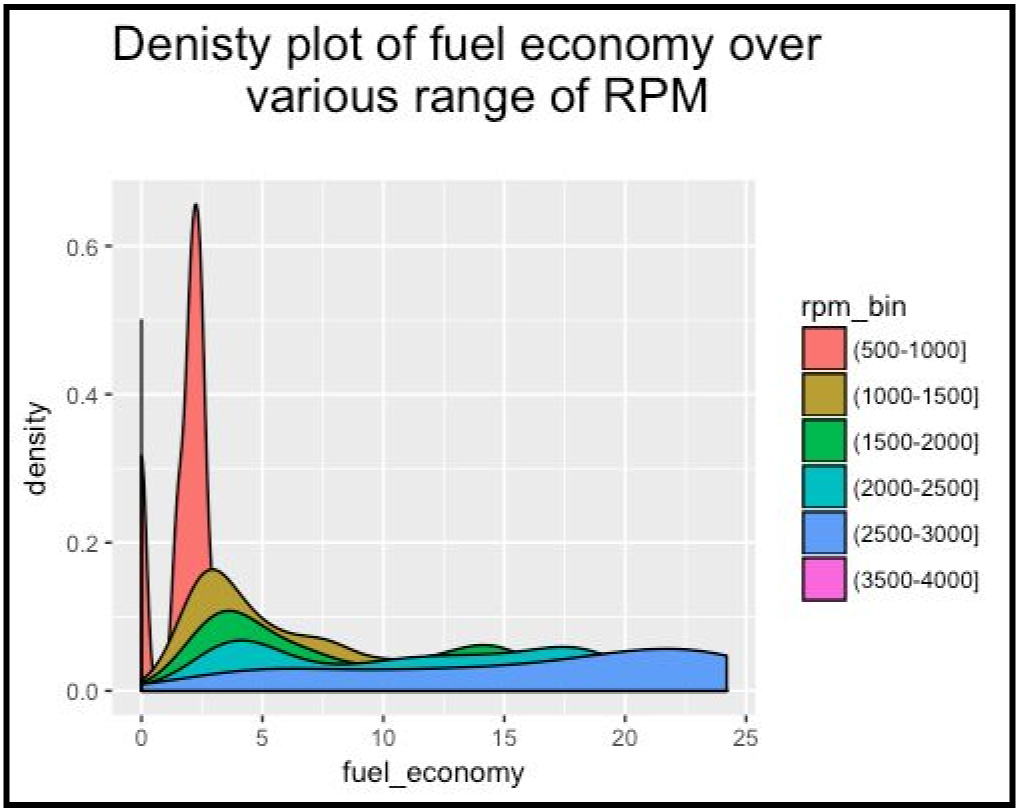


Explanation:

It explains the density of fuel_economy for different range of speeds.

Conclusion:

Speed range 60-70 has produced mileage from 0 to 23



Explanation:This graph explains the density of fuel_economy over different RPM's

Suggestions:

1. Time of the instance in not in periodic.It is collecting randomly.

   We can draw some other conclusions if it is periodic.

   Rest all is good for a garage owner or car owner to identify how the car was driven and what are the things to do based on the track data.

Chapter 4

Code:

```
library(readxl)
orig_data= read_excel("Shanrohi.xlsx")[c(-1,-1445),]
colnames(orig_data)=c('date','time','speed','fuel_economy','rpm','coolant_temp')

#install.packages('lubricate')
library(lubridate)
orig_data$time <- factor(orig_data$time)

# parese date
a <- hms(as.character(orig_data$time))

# get hours
orig_data$hour=hour(a)

# get minutes
minute(a)
switch_on=data.frame(orig_data[orig_data$speed==0 & orig_data$rpm==0,])
#data[is.na(data)]=0
#selecting the rows with speed=0 but rpm != 0
Stable_data=orig_data[orig_data$speed==0 & orig_data$rpm !=0,]

data=orig_data[,3:7]
data[data ==0] <- NA
#data[which(is.na(data)),]
data=data[apply(data[,1:4],1,function(x)any(!is.na(x))),]
data[is.na(data)] <- 0
```

```
#new_DF <- data[rowSums(is.na(data[,3:6])) > 0,]
#a=data.frame(data[is.na(data[,3:6]),])
#data=na.omit(data)
#install.packages('stringr')
library(stringr)
data$speed=gsub(' N',"",data$speed)
data$coolant_temp=gsub("F","",data$coolant_temp)
data[,1:4]=lapply(data[,1:4],as.numeric)
#hist(data$fuel_economy)
data[data$fuel_economy==242320.4 & data$speed==15776,]=NA
data=na.omit(data)
data$speed_bin=cut(data$speed, breaks = c(-Inf,0,10,20,30,40,50,60,70), labels =
c('Stable','(0-10]','(10-20]','(20-30]',
                                                  '(30-40]','(40-50]','(50-60]','(60-70]'))



data$rpm_bin=cut(data$rpm,breaks=c(500,1000,1500,2000,2500,3000,3500,4000),labels=

c('(500-1000]','(1000-1500]','(1500-2000]','(2000-2500]','(2500-3000]','(3000-3500]','(3500-400
0]'))


data$economy=factor(ifelse(data$fuel_economy<17,'less than 17','Greater than 17' ))
data$hour=as.factor(data$hour)


library(ggplot2)
ggplot(data,aes(fuel_economy))+geom_histogram(aes(fill=..count..),bins=10)
#data$fuel_economy
#data=na.omit(data)
#data=na.omit(data)
```

```
sum(is.na(data))
#density plot of fuel_economy for RPM and Speed ranges
ggplot(data,aes(x=fuel_economy,y=..density..,fill=speed_bin))+geom_density()+
  ggtitle("Denisty plot of fuel economy over
          various range of speed\n") +
  theme(plot.title=element_text(size=18))
ggplot(data,aes(x=fuel_economy,y=..density..,fill=rpm_bin))+geom_density()+
  ggtitle("Denisty plot of fuel economy over
          various range of RPM\n") +
  theme(plot.title=element_text(size=18))
ggplot(data,aes(x=coolant_temp))+geom_histogram(aes(fill=..count..),bins=70)+
  ggtitle("Histogram of coolant temp\n") +
  theme(plot.title=element_text(size=18))
ggplot(data,aes(x=coolant_temp,y=..density..,fill=speed_bin))+geom_density()
ggplot(data,aes(x=coolant_temp,y=..density..,fill=rpm_bin))+geom_density()



#rpm vs fuel economy
ggplot(data,aes(x=rpm_bin,y=fuel_economy,colour=rpm_bin))+geom_point()+

annotate('rect',ymin=17,ymax=23,xmax='(500-1000]',xmin='(3500-4000]',alpha=0.3,fill='green')+
  annotate('text',y=24,x='(2000-2500]',label="Range of RPM's which gave the ideal mileage")+geom_hline(yintercept = 17.4)



# 3D Exploded Pie Chart
#install.packages('plotrix')
library(plotrix)
```

```
slices = c(113,111,223,329,34,2)
lbls=c('(500-1000]','(1000-1500]','(1500-2000]','(2000-2500]','(2500,3000]','(3500-4000]')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Pie Chart of RPM of the car(Total data) ")

#performance of car(mileage) over different RPM ranges
ggplot(data,aes(x=rpm_bin,y=fuel_economy,colour=rpm_bin))+geom_point(shape=20)+
  geom_hline(yintercept = 17.4)+
  annotate('text',y=24,x='(2000-2500]',label='Ideal mileage is
           obtained only in 2500-4000 Rpm ranges')

#speed of the car over different RPM ranges
ggplot(data,aes(x=rpm_bin,y=speed,colour=rpm_bin))+geom_point()+
  ggtitle("variation in speed of the car driven over different RPM ranges.\n") +
  theme(plot.title=element_text(size=18))

#  how the car has performed(mileage) with RPM over the speed ranges
ggplot(data,aes(x=rpm_bin,y=fuel_economy,colour=speed_bin))+geom_point(shape=20)+
  geom_hline(yintercept = 17.4)+ggtitle("RPM vs Fuel_economy with different speed
ranges\n") +
  theme(plot.title=element_text(size=18))

#fuel economy between 17 to 23 in findinf ideal RPM
fuel_economy_data=data[data$fuel_economy>=17 & data$fuel_economy<23,]
slices = c(1,71,11)
lbls=c('(1500-2000]-','(2000-2500]-','(2500-3000]-')
pct=round(slices/sum(slices)*100)
```

```
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main=" RPM of the car driven which resulted
    mileage bewtween 17 and 23 in total data")

# data with rpm(2000-2500] and mileage between 17 and 23
economy_rpm_fuel=fuel_economy_data[fuel_economy_data$rpm_bin=='(2000-2500]' |
fuel_economy_data$rpm_bin=='(2500-3000]',]

slices = c(1,3,5,51,22)
lbls=c('(10-20)-','(20-30)-','(30-40)-','(40-50)-','(50-60)-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="range of speed of the car driven Which resulted in
    mileage between 17 and 23 under rpm of 2000-3000]")

############
#
#
# Analysing the RPM where the car has not performed well
#
#
##########



#data for which the fuel_economy is less than 17
```

```
less_economy=data[data$fuel_economy<17,]
slices = c(113,111,222,258,14,2)
lbls=c('(500_1000]-','(1000-1500]-','(1500-2000]-','(2000-2500]-','(2500-3000]-','(3500-4000]')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Pie Chart of RPM of the car which resulted in
    mileage less than 17 in total data")


less_economy_speed=less_economy[less_economy$rpm>1000 &less_economy$rpm<2500,]
table(less_economy_speed$speed_bin)
slices = c(8,14,62,78,141,224,63)
lbls=c('Stable','(0-10]-','(10-20)-','(20-30)-','(30-40)-','(40-50)-','(50-60)-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Pie Chart of speed of the car which resulted in
    mileage less than in 17  under RPM of 1000-2500 in total data")

ggplot(data,aes(x=rpm,y=speed,colour=economy))+geom_point()+
  annotate('rect',xmin=2000,xmax=3000,ymin=30,ymax=60,alpha=0.3,fill='green')+
  annotate('text',x=3000,y=30,label='Region where fuel_economy
         is varying for same rpm and speed',color="blue")+
  ggtitle("speed vs rpm  with respect to fuel_economy\n") +
  theme(plot.title=element_text(size=18))+
  annotate('rect',xmin=500,xmax=2000,ymin=0,ymax=45,alpha=0.3,fill='red')+
```

```
  annotate('text',x=2500,y=10,label='Region where fuel_economy
        is less for this ranges of rpm and speed',color="blue")
```

```
###########################
#
#
#speed vs fuel economy
#
#
##########################
```

```
#speed vs fuel economy
ggplot(data,aes(x=speed_bin,y=fuel_economy,colour=speed_bin))+geom_point()+
  annotate('rect',ymin=17,ymax=23,xmax='Stable',xmin='(60-70]',alpha=0.3,fill='green')+
  annotate('text',y=24,x='(20-30]',label="Region explaining
        Range of speeds which gave the ideal mileage")+geom_hline(yintercept = 17.4)
```

```
#pie chaart of speeds of the car driven
table(data$speed_bin)
slices = c(69,39,79,89,148,277,108,3)
```

```
lbls=c('Stable','(0-10)-','(10-20)-','(20-30)-','(30-40)-','(40-50)-','(50-60)-','(60-70)-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Pie Chart of speed of the car driven
    in various speed ranges")



#performance of car(mileage) over different speed ranges
ggplot(data,aes(x=speed_bin,y=fuel_economy,colour=speed_bin))+geom_point(shape=20)+
  geom_hline(yintercept = 17.4)+
  annotate('text',y=24,x='30',label='Ideal mileage is
          obtained only in  speed ranging from 30-60')

# RPM of the car over different speed ranges
ggplot(data,aes(x=speed_bin,y=rpm,colour=speed_bin))+geom_point()+
  ggtitle("variation in rpm of the car driven over
         different speed ranges.\n") +
  theme(plot.title=element_text(size=18))

#  how the car has performed(mileage) with RPM over the speed ranges
ggplot(data,aes(x=speed_bin,y=fuel_economy,colour=rpm_bin))+geom_point(shape=20)+
  geom_hline(yintercept = 17.4)+ggtitle("Speed vs fuel economy for different
                                     RPM ranges\n") +
  theme(plot.title=element_text(size=18))
```

```
# pie chart of speed with obtained mileage between 17 and 23
table(fuel_economy_data$speed_bin)
slices = c(1,3,5,52,22)
lbls=c('(10-20)-','(20-30)-','(30-40)-','(40-50)-','(50-60)-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Speed of the car driven for obtaining
    mileage bewtween 17 and 23 in total data")



#data with speed=40-60 and mileage between 17 and 23table(less_economy$speed_bin)
economy_speed_fuel=fuel_economy_data[fuel_economy_data$speed_bin==50 |
                          fuel_economy_data$speed_bin==60,]
table(economy_speed_fuel$rpm_bin)
slices = c(1,63,10)
lbls=c('(1500-2000]-','(2000-2500]-','(2500-3000]-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="RPM of the car driven  with in speed 40- 60 with obtained
    mileage between 17 and 23 \n.")

########################
#
```

```
#
# Analysing the speed where the car has not performed well
#
#
################################
table(less_economy$speed_bin)

slices = c(69,39,78,86,143,225,79,1)
lbls=c('Stable-','(0-10]-','(10-20]-','(20-30]-','(30-40]-','(40-50]-','(50-60]-','(60-70]-')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="speed of the car driven  with obtained mileage less than 17")

#pints where mileage will be min
ggplot(less_economy,aes(x=rpm_bin,y=speed,colour=economy))+geom_point()+
  ggtitle("Speed vs rpm  for fuel_economy <17 \n")

ggplot(fuel_economy_data,aes(x=rpm,y=speed))+geom_point()

slices = c(113,111,222,258,14,2)
lbls=c('(500_1000]-','(1000-1500]-','(1500-2000]-','(2000-2500]-','(2500-3000]-','(3500-4000]')
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct) # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
pie3D(slices,labels=lbls,explode=0.1,
    main="Pie Chart of RPM of the car for obtaining
    mileage less than in 17 in total data")
```

```
#combination of speed and rpm to get the mileage greater than 17
ggplot(data,aes(x=rpm,y=speed,colour=economy))+geom_point()+
  annotate('rect',xmin=2000,xmax=3000,ymin=30,ymax=60,alpha=0.3,fill='green')+
  annotate('text',x=3000,y=30,label=' Green Region where fuel_economy
           is varying(>17 and <17) for same rpm and speed',color="blue")+
  ggtitle("speed vs rpm  with respect to fuel_economy\n") +
  theme(plot.title=element_text(size=18))+
  annotate('rect',xmin=500,xmax=2000,ymin=0,ymax=45,alpha=0.3,fill='red')+
  annotate('text',x=2500,y=10,label=' Red Region where fuel_economy
           is less(<17) for this ranges of rpm and speed',color="blue")


#Explaning the car being brough to neutral while driving
Stable_data$hour=as.factor(Stable_data$hour)
Stable_data$fuel_economy=as.numeric(Stable_data$fuel_economy)
ggplot(Stable_data,aes(x=hour,colour=hour))+geom_bar()+
  ggtitle("Number of times car brought to stable position(Speed=0,RPM !=0 )
           after some speed has attained in different hours\n")

#effect of mileage for being to to mileage in different hours
ggplot(Stable_data,aes(x=hour,y=fuel_economy,colour=hour))+geom_point()+
  ggtitle("fuel_economy of the car in different hours when car was
          brought to stable(speed=0,RPM !=0)  \n")+ ylim(1,23)
# % of the time that engine was off
slices=c(29,26,27,24,26,24,39,36,37,46,47,28,26,37,50,45,34,24,25 )
lbls=c(0:18)
pct=round(slices/sum(slices)*100)
lbls=paste(lbls, pct,sep="hr -") # add percents to labels
lbls=paste(lbls,"%",sep="") # ad % to labels
```

```
pie3D(slices,labels=lbls,explode=0.1,
      main="percentage of time of car where
      speed=0=rpm (key on,engine off)")
switch_on$hour=as.factor(switch_on$hour)
ggplot(switch_on,aes(x=hour,fill=hour))+geom_bar()+
  annotate('rect',xmin=7,xmax=)
```