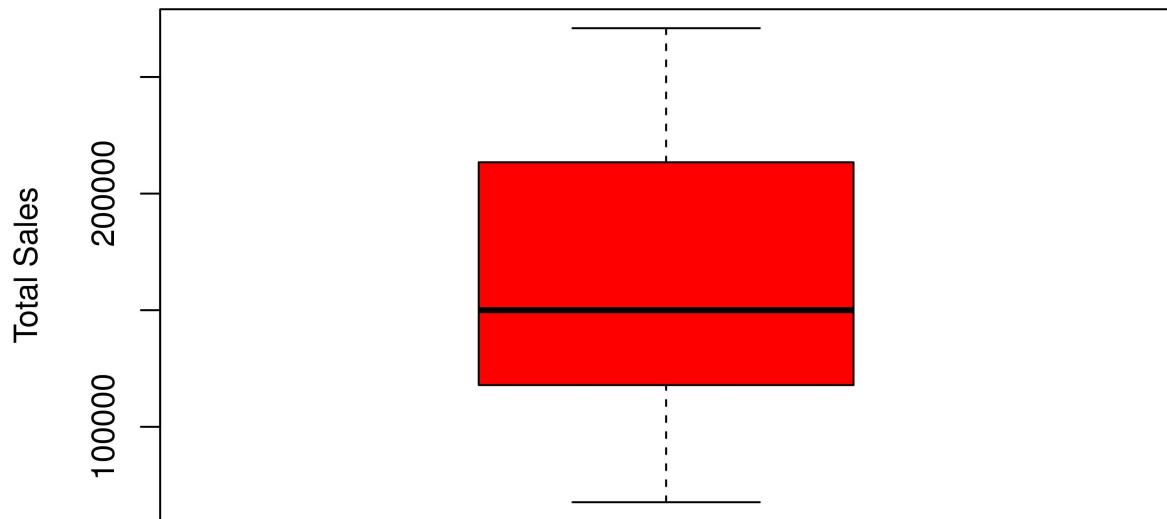


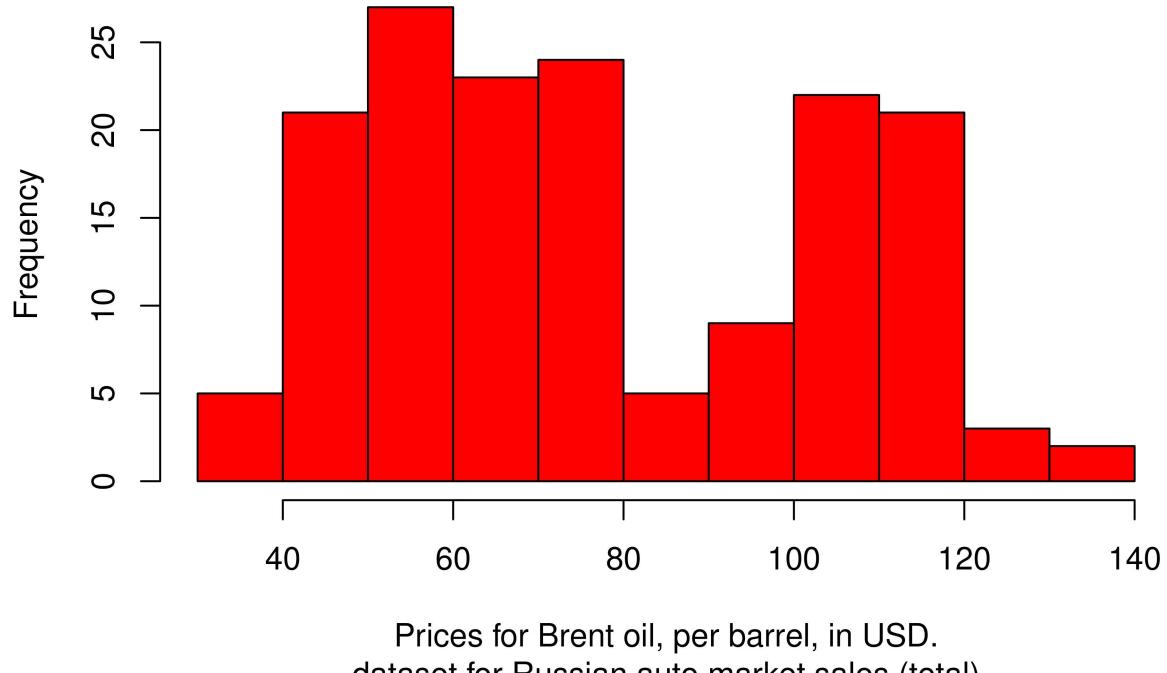
Boxplot for Total Sale (from jan 2005 to jun 2018)



dataset for Russian auto market sales (total), starting from the year 2005 to 2018

```
# Histogram for oil-
oil <- auto_data$Oil
hist(oil, xlab="Prices for Brent oil, per barrel, in USD.", sub= "dataset for Russian auto market sa")
```

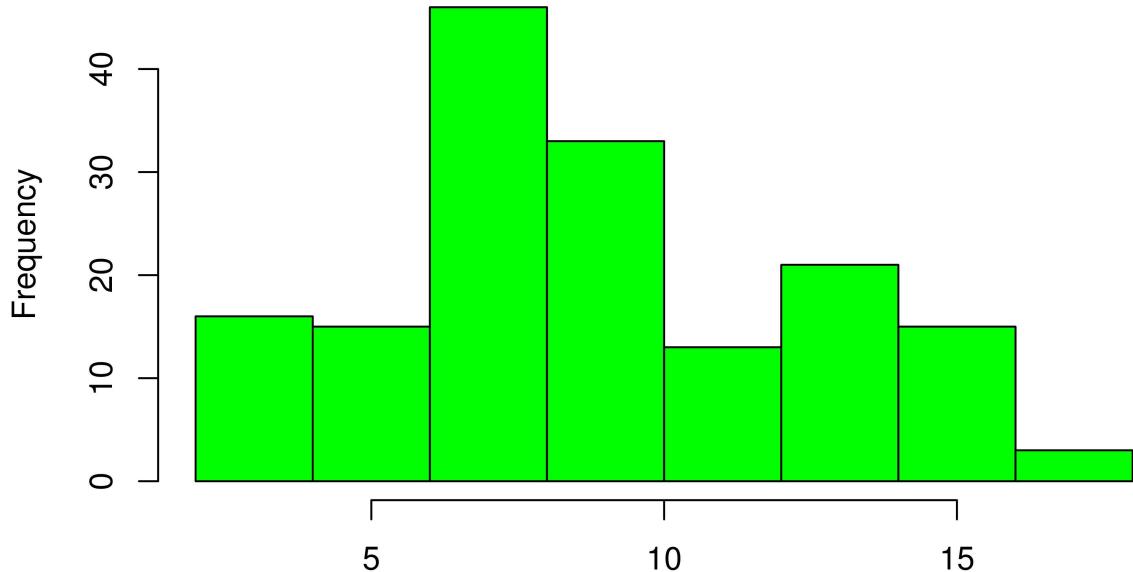
Histogram for oil price from january 2005 to june 2018



Histogram for CPI:-

```
cpi <- auto_data$CPI  
hist(cpi, xlab="Consumer price index CPI(inflation)", sub= "dataset for Russian auto market sales (t")
```

Histogram for Consumer Price Index(CPI) from january 2005 to june 2017



pair plot in auto_data

```
if(!require('GGally'))  
{  
  install.packages('GGally')  
  library(GGally)  
}  
  
## Loading required package: GGally  
  
## Warning: package 'GGally' was built under R version 4.1.1  
  
## Registered S3 method overwritten by 'GGally':  
##   method from  
##   +.gg   ggplot2
```

4. For the same variables, calculate the appropriate measures of central tendency.

```
if(!require('dplyr'))  
{  
  install.packages('dplyr')  
  library(dplyr)  
}
```

TotalSales	Oil	CPI
Min. : 67670	Min. : 30.70	Min. : 2.200
1st Qu.:118181	1st Qu.: 55.95	1st Qu.: 6.450
Median :150011	Median : 72.31	Median : 8.110
Mean :161963	Mean : 77.48	Mean : 8.859
3rd Qu.:213279	3rd Qu.:106.01	3rd Qu.:11.822
Max. :270914	Max. :132.72	Max. :16.930

	TotalSales	Oil	CPI
TotalSales	1.000	0.786	-0.094
Oil	0.786	1.000	-0.196
CPI	-0.094	-0.196	1.000

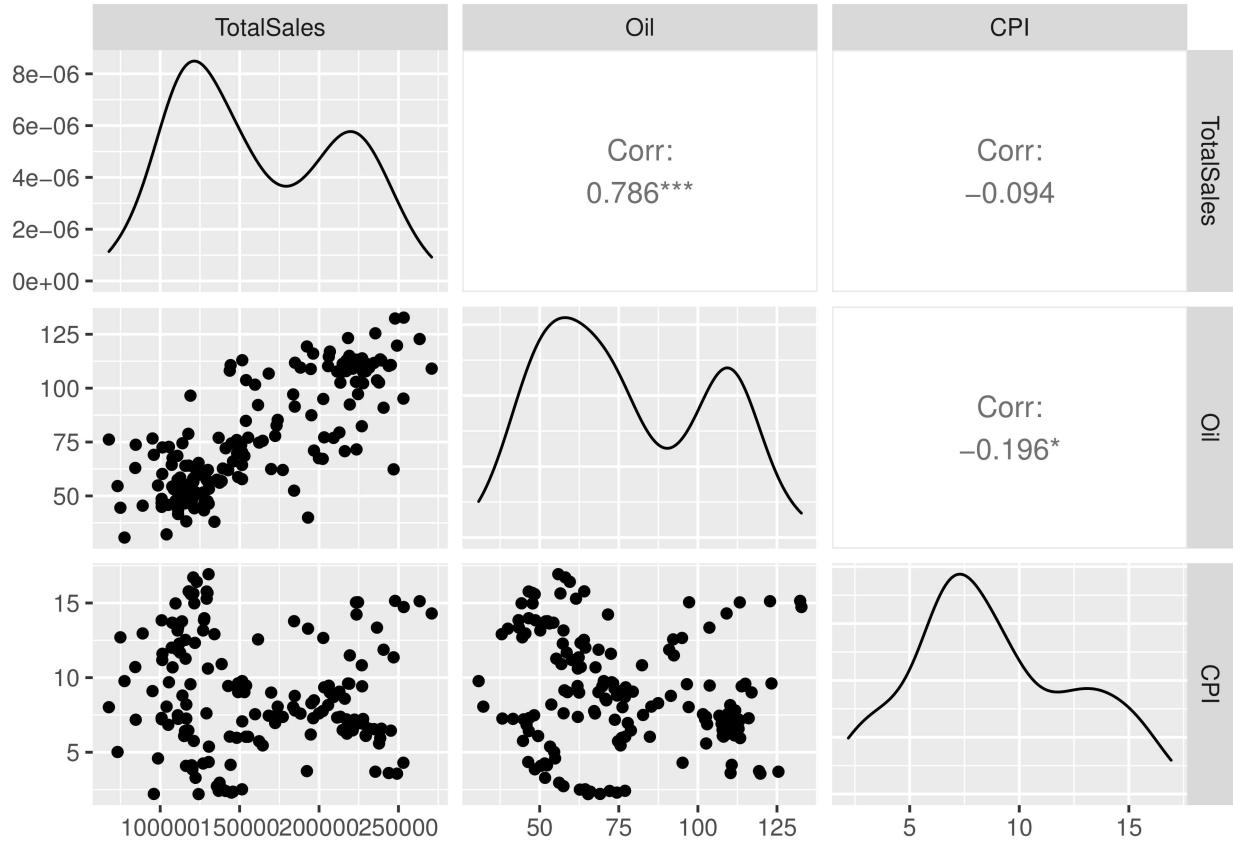
```
ct <- auto_data %>%
  select("TotalSales","Oil","CPI")
sum_ct <- summary(ct)
kable(sum_ct, booktabs= TRUE) %>%
  kable_styling()
```

- For the automarket “Total Sales” a monthly mean of 161963 sales per month can be observed with a minimum of 67670 sales and maximum of 270914 sales.
- For the variable oil mean of 77.48 USD can be observed with a min of 30.70 and max of 132.72.
- For the variable cpi mean of 8.859% can be observed with a min of 2.200 and max of 16.930.

5. For the same variables, create a matrix of correlations.

```
corrmat <- cor(ct,method = "pearson",use = "complete.obs")
kable(corrmat,booktabs=TRUE, digits = 3) %>%
  kable_styling()
```

```
comat <- auto_data %>%
  select("TotalSales", "Oil","CPI")
ggpairs(comat)
```



- The strength and direction of linear relationship between two variables can be described with the correlation coefficient. A perfect correlation has a value of +1 (positive correlation) and -1 (negative correlation).
 - For the chosen variables following correlation coefficients are observed.
 - * Total Sales and Oil: - 0.786 which is a strong positive correlation.
 - * Total Sales and CPI: - 0.094 which is a weak negative correlation.
 - * Oil and CPI : - 0.196 which is a weak negative correlation.

6. Create at least one complex chart that combines information.

```
if(!require('dplyr'))
{
  install.packages('dplyr')
library(dplyr)
}

if(!require('tibble'))
{
  install.packages('tibble')
library(tibble)
}
```

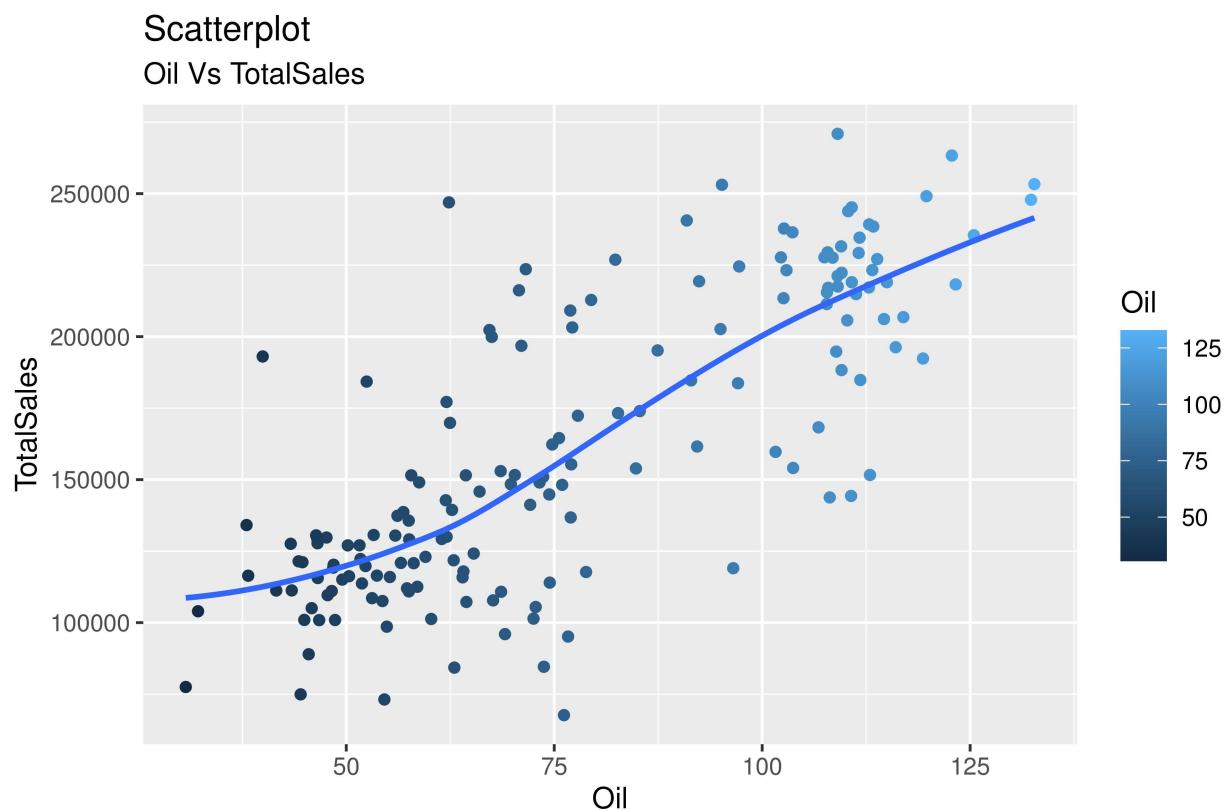
```
# Scatterplot
```

```

gg <- ggplot(auto_data, aes(x=Oil, y=TotalSales)) +
  geom_point(aes(col=Oil)) +
  geom_smooth(method="loess", se=F) +
  labs(subtitle="Oil Vs TotalSales",
       y="TotalSales",
       x="Oil",
       title="Scatterplot",
       caption = "Source: dataset for Russian auto market sales (total), starts from january 2005")
plot(gg)

```

'geom_smooth()' using formula 'y ~ x'



Source: dataset for Russian auto market sales (total), starts from january 2005 to june 2018

7. Create at least one complex chart that demonstrates results of the model. To do so, perform the following tasks:

- Choose a dependent variable
- Choose at least two independent variables
- Build a linear regression model
- Plot results of the model

```

if(!require('corrplot'))
{
install.packages('corrplot')

```

	TotalSales	CPI	Oil	USD	CCI	BCI	Primerate
TotalSales	1.00	-0.09	0.79	-0.42	0.55	0.01	-0.40
CPI	-0.09	1.00	-0.20	-0.20	-0.33	-0.44	0.72
Oil	0.79	-0.20	1.00	-0.51	0.58	0.10	-0.61
USD	-0.42	-0.20	-0.51	1.00	-0.59	0.19	0.06
CCI	0.55	-0.33	0.58	-0.59	1.00	0.50	-0.40
BCI	0.01	-0.44	0.10	0.19	0.50	1.00	-0.31
Primerate	-0.40	0.72	-0.61	0.06	-0.40	-0.31	1.00

```
library(corrplot)
}

## Loading required package: corrplot

## Warning: package 'corrplot' was built under R version 4.1.1

## corrplot 0.90 loaded
```

remove date from the dataset which is not usable in the correlation:

```
autodata_model <- auto_data[,-1]
```

correlations of all other variables in the auto_data dataset:

```
corrModel <- cor(autodata_model, method = "pearson", use = "complete.obs")

kable(corrModel, booktabs=TRUE, digits = 2) %>%
  kable_styling()
```

```
pairs(corrModel)
```