Regression Analysis

MANOJ KUMAR - 2048015

30/01/2021

## 1.Import the dataset data\_marketing\_budget\_mo12 and do the exploratory data analysis .

# Importing the shared Marketing Budget Data.  
  
RocketData <- read.csv("data-marketing-budget.csv")  
RocketData

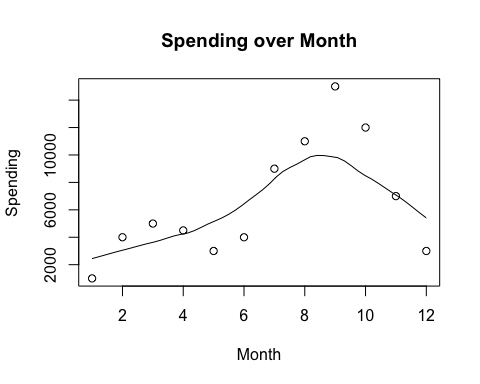
## Month Spend Sales  
## 1 1 1000 9914  
## 2 2 4000 40487  
## 3 3 5000 54324  
## 4 4 4500 50044  
## 5 5 3000 34719  
## 6 6 4000 42551  
## 7 7 9000 94871  
## 8 8 11000 118914  
## 9 9 15000 158484  
## 10 10 12000 131348  
## 11 11 7000 78504  
## 12 12 3000 36284

# Summary  
summary(RocketData)

## Month Spend Sales   
## Min. : 1.00 Min. : 1000 Min. : 9914   
## 1st Qu.: 3.75 1st Qu.: 3750 1st Qu.: 39436   
## Median : 6.50 Median : 4750 Median : 52184   
## Mean : 6.50 Mean : 6542 Mean : 70870   
## 3rd Qu.: 9.25 3rd Qu.: 9500 3rd Qu.:100882   
## Max. :12.00 Max. :15000 Max. :158484

## 2. Use Scatter Plot To Visualise The Relationship.

scatter.smooth(x = RocketData$Month,   
 y = RocketData$Spend,   
 xlab = "Month", ylab = "Spending", # X,Y-Lab  
 main = "Spending over Month", # Title  
 col.main = "Black", # Title lab color  
 cex.lab = "1" # X,Y-axis Lab color  
)

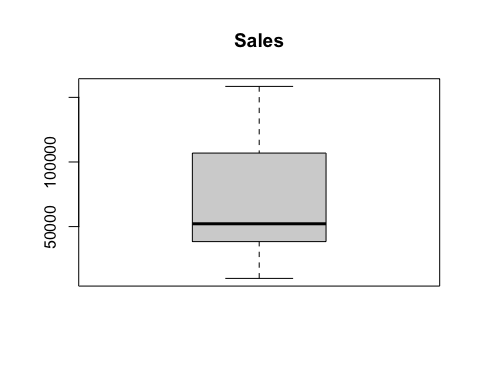


scatter.smooth(x = RocketData$Month,   
 y = RocketData$Sales,   
 xlab = "Month", ylab = "Sales", # X,Y-Lab  
 main = "Sales over Month", # Title  
 col.main = "Black", # Title lab color  
 cex.lab = "1" # X,Y-axis Lab color   
)

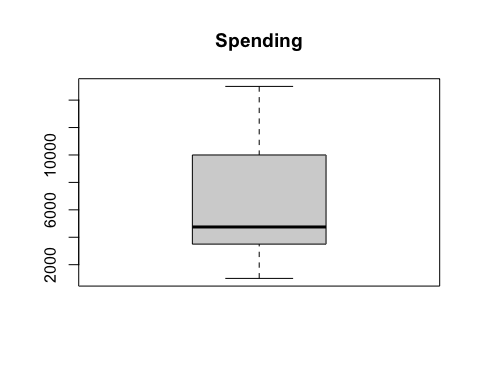


## 3.Using BoxPlot To Check For Outliers.

boxplot(RocketData$Sales, main = "Sales" )

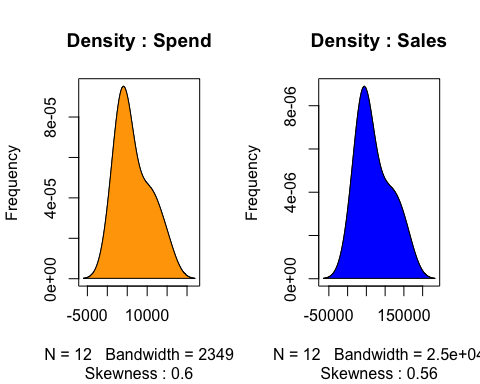


boxplot.default(RocketData$Spend, main = "Spending" )



## 4. Using Density Plot To Check If Response Variable Is Close To Normal.

library(e1071) # For skewness function  
par(mfrow = c(1,2)) # Dividing graph area in 2 columns  
  
# Density plot for 'Spend'  
plot(density(RocketData$Spend),   
 main = "Density : Spend",  
 ylab = "Frequency",  
 sub=paste("Skewness :",  
 round(e1071::skewness(RocketData$Spend),  
 2  
 )  
 )  
 )   
polygon(density(RocketData$Spend), col = "orange")  
  
# Density plot for 'Sales'  
plot(density(RocketData$Sales),   
 main = "Density : Sales",  
 ylab = "Frequency",  
 sub=paste("Skewness :",  
 round(e1071::skewness(RocketData$Sales),  
 2  
 )  
 )  
 )   
  
polygon(density(RocketData$Sales), col = "blue")



## 5. Check the Correlation Analysis.

cor(RocketData$Spend, RocketData$Sales)

## [1] 0.9988322

## 6. Build the Linear Regression Model.

LinearRegressionModel <- lm(formula = Spend ~ Sales, data = RocketData)  
LinearRegressionModel

##   
## Call:  
## lm(formula = Spend ~ Sales, data = RocketData)  
##   
## Coefficients:  
## (Intercept) Sales   
## -114.67027 0.09392

summary(LinearRegressionModel)

##   
## Call:  
## lm(formula = Spend ~ Sales, data = RocketData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -293.22 -165.15 -20.82 188.67 312.02   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.147e+02 1.196e+02 -0.959 0.36   
## Sales 9.392e-02 1.437e-03 65.378 1.71e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 217.5 on 10 degrees of freedom  
## Multiple R-squared: 0.9977, Adjusted R-squared: 0.9974   
## F-statistic: 4274 on 1 and 10 DF, p-value: 1.707e-14

## 7. Using p-value Check For Statistical Significance.

## 8.Capture the summary of the linear model.

ModelSummary <-summary(LinearRegressionModel)  
ModelSummary

##   
## Call:  
## lm(formula = Spend ~ Sales, data = RocketData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -293.22 -165.15 -20.82 188.67 312.02   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.147e+02 1.196e+02 -0.959 0.36   
## Sales 9.392e-02 1.437e-03 65.378 1.71e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 217.5 on 10 degrees of freedom  
## Multiple R-squared: 0.9977, Adjusted R-squared: 0.9974   
## F-statistic: 4274 on 1 and 10 DF, p-value: 1.707e-14

## 9. Also perform the Linear Diagnostics for the given data set.(Hint: plot(lmmodel) )

plot(LinearRegressionModel)

