#LOAD LIBRARY

library(lattice)

library(psych)

library(gmodels)

library(ggplot2)

library(dplyr)

#Describe dataset

describeData(MultiRegDataset)

#Summary of dataset

summary(MultiRegDataset)

#Histogram

dependent\_var = MultiRegDataset$expenses

# histogram with added parameters

hist(MultiRegDataset$expenses,

main="Histogram of our Dependent Variable - Expenses",

xlab="Expenses",

col="beige",

)

#One Sample t-test = 15000

t.test(MultiRegDataset$expenses, mu=15000)

ggplot(MultiRegDataset, aes(x = expenses, y = smoker))+

geom\_point(colour = "red") + geom\_smooth(method = "lm", fill = NA) +

labs(title = "Expenses vs Smoker",

x = "Expenses",

y = "Smoker")

#Correlation & ENCoder

#Encode Smoker Column from yes/no to 0/1)

MultiRegDataset$smoker <- ifelse(MultiRegDataset$smoker=="yes",1,0)

MultiRegDataset

#Correlation

cor(MultiRegDataset$smoker,MultiRegDataset$expenses)

#Linear Model - No Transformation

lmodel <- lm(MultiRegDataset$expenses ~ MultiRegDataset$smoker, data = MultiRegDataset)

summary(lmodel)

#Multiple Regression Model

#Create Model2

model2 <- lm(expenses~., data = MultiRegDataset)

#Summary of Model2

print(model2)

summary(model2)