29th Jan Data : Usedcarprice

**Multiple linear regression class**

usedcarprice=read.csv(file.choose())

head(usedcarprice)

tail(usedcarprice)

str(usedcarprice)

summary(usedcarprice)

summary(usedcarprice$Price)

class(usedcarprice)

nrow(usedcarprice)

ncol(usedcarprice)

hist(usedcarprice$Price)

plot(hist(log(usedcarprice$Price)))

boxplot(usedcarprice$Price,horizontal = T)

plot(density(usedcarprice$Price))

plot(density(log(usedcarprice$Price)))

table(usedcarprice$FuelType)

table(usedcarprice$AutoType)

str(usedcarprice)

aggregate(usedcarprice$Price~usedcarprice$AutoType,FUN = mean)

pairs(usedcarprice)

t.test(usedcarprice$Price~usedcarprice$Automatic)

aggregate(usedcarprice$Price~usedcarprice$FuelType, FUN = mean)

summary(aov(usedcarprice$Price~usedcarprice$FuelType))

summary(usedcarprice)

usedcarprice$ageclass=ifelse(usedcarprice$Age>=70, "very old",ifelse(usedcarprice$Age>=55,"old",ifelse(usedcarprice$Age>=44,"recent","new")))

summary(usedcarprice$ageclass)

table(usedcarprice$ageclass)

table(usedcarprice$ageclass,usedcarprice$FuelType)

chisq.test(table(usedcarprice$ageclass,usedcarprice$FuelType))

# Simple Multiple Linear Regression example

str(usedcarprice)

usedcarprice1=usedcarprice[-c(1,7,8)]

usedcarnumeric=Filter(is.numeric,usedcarprice1) # Filter command will select only the type of variables(categorical, numeric orinterger using is.\_\_)

str(usedcarnumeric)

cov(usedcarnumeric)

pairs(usedcarnumeric)

usedcarreg1=lm(Price~.,data=usedcarprice1)

summary(usedcarreg1)

colnames(usedcarprice1)

usedcarreg2=lm(Price~Age+KM+FuelType+HP+CC+Doors+Weight+AutoType+MetColorType+ageclass,data=usedcarprice1)

summary(usedcarreg2)

anova(usedcarreg1,usedcarreg2)

plot(usedcarreg2)

usedcarreg2$residuals # errors

usedcarreg2$fitted.values # Predicted value

sqrt(mean(usedcarreg2$residuals^2)) #RMSE

#step selection of variables in reg model

usedcarstep=step(usedcarreg1)

summary(usedcarstep)

write.csv(usedcarstep,"usedcarreg2.csv")

cars=mtcars

write.csv(mtcars,"mtcars.csv")

getwd() # will return theworking file directory location