Class5\_script.R

daniellee

Wed Feb 15 21:10:13 2017

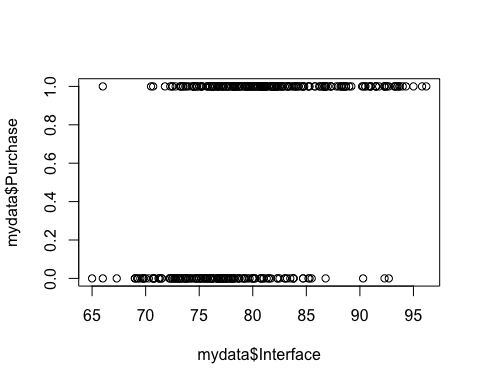
mydata = read.csv("/Users/daniellee/Desktop/BC/Study/Class Resource/2nd/Marketing Infro. Anal./Class5/App.csv")  
  
?glm  
#Check the data  
str(mydata)

## 'data.frame': 394 obs. of 3 variables:  
## $ Interface: num 78 73 72.3 72.8 73.8 83.5 75.5 74.5 69 75.5 ...  
## $ LoadTime : int 80 77 60 74 60 67 78 65 62 72 ...  
## $ Purchase : int 0 0 0 0 0 0 0 0 0 0 ...

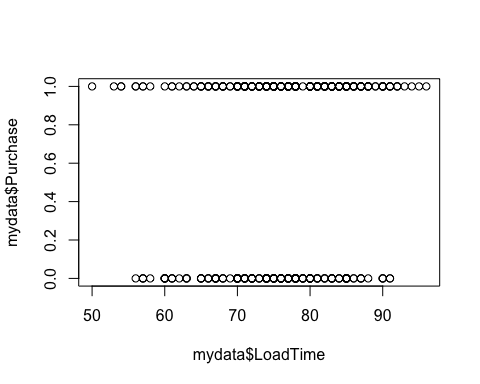
summary(mydata)

## Interface LoadTime Purchase   
## Min. :65.00 Min. :50.00 Min. :0.000   
## 1st Qu.:75.70 1st Qu.:71.00 1st Qu.:0.000   
## Median :78.70 Median :77.00 Median :1.000   
## Mean :79.74 Mean :76.29 Mean :0.599   
## 3rd Qu.:83.00 3rd Qu.:83.00 3rd Qu.:1.000   
## Max. :96.20 Max. :96.00 Max. :1.000

plot(mydata$Interface, mydata$Purchase)



plot(mydata$LoadTime, mydata$Purchase)



#Run the logistic model  
model1 = glm(Purchase ~ Interface, data = mydata, family = "binomial")  
  
summary(model1) #Linear affect on the Z value

##   
## Call:  
## glm(formula = Purchase ~ Interface, family = "binomial", data = mydata)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.5152 -1.0209 0.4435 0.9321 2.1302   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -15.21013 1.99832 -7.611 2.71e-14 \*\*\*  
## Interface 0.19773 0.02548 7.759 8.56e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 530.66 on 393 degrees of freedom  
## Residual deviance: 443.43 on 392 degrees of freedom  
## AIC: 447.43  
##   
## Number of Fisher Scoring iterations: 4

exp(coef(model1)) #Calculate the odds ratio

## (Intercept) Interface   
## 2.479276e-07 1.218639e+00

#For an increase of 1 unit in the interface, the chance of buying incerease 1.22  
  
#Compare model1  
model2 = glm(Purchase ~ Interface + LoadTime, data = mydata, family = "binomial")  
summary(model2) #Linear affect on the Z value

##   
## Call:  
## glm(formula = Purchase ~ Interface + LoadTime, family = "binomial",   
## data = mydata)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.5577 -0.9833 0.4340 0.9126 2.2883   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -14.69568 2.00683 -7.323 2.43e-13 \*\*\*  
## Interface 0.22982 0.02955 7.776 7.47e-15 \*\*\*  
## LoadTime -0.04020 0.01709 -2.352 0.0187 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 530.66 on 393 degrees of freedom  
## Residual deviance: 437.69 on 391 degrees of freedom  
## AIC: 443.69  
##   
## Number of Fisher Scoring iterations: 4

#Positie on Interface and Negatice on LoadTime  
#AIC is lower, better  
exp(coef(model2))

## (Intercept) Interface LoadTime   
## 4.147134e-07 1.258378e+00 9.605967e-01

#9.605967e-01 smaller than one, the chance of buying decrease  
  
#Max  
logLik(model1)

## 'log Lik.' -221.7174 (df=2)

logLik(model2)

## 'log Lik.' -218.8427 (df=3)