> auto<- read.csv(file.choose(), header=TRUE)

> attach(auto)

> cy=(CLMINSUR==1)

> sy=(SEATBELT==1)

> model1=lm(LOSS~cy+sy+ATTORNEY+CLMSEX+factor(MARITAL)+CLMAGE)

> summary(model1)

Call:

lm(formula = LOSS ~ cy + sy + ATTORNEY + CLMSEX + factor(MARITAL) +

CLMAGE)

Residuals:

Min 1Q Median 3Q Max

-19.903 -4.518 -1.846 0.466 264.106

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 29.33265 4.23368 6.928 7.29e-12 \*\*\*

cyTRUE -2.00846 1.62472 -1.236 0.217

syTRUE -14.59133 3.55391 -4.106 4.33e-05 \*\*\*

ATTORNEY -6.39352 0.96216 -6.645 4.80e-11 \*\*\*

CLMSEX -0.38066 0.96940 -0.393 0.695

factor(MARITAL)2 -1.66404 1.14918 -1.448 0.148

factor(MARITAL)3 -2.81182 4.72568 -0.595 0.552

factor(MARITAL)4 0.69435 3.00151 0.231 0.817

CLMAGE 0.04156 0.03410 1.219 0.223

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.7 on 1082 degrees of freedom

(249 observations deleted due to missingness)

Multiple R-squared: 0.06457, Adjusted R-squared: 0.05765

F-statistic: 9.336 on 8 and 1082 DF, p-value: 1.619e-12

a) #Only syTRUE IS significant,since its pvalue = 4.33e-05<.05(alpha)

#while cyTRUE is not,since .217 >0.05

b) Its R sqr and R sqr adj are both very low,so a lot of omitting data affects this model’s prediction

yet it has a very low Fp-value <.05

> 29.33265-.38066

[1] 28.95199

c)## as cyFalse and syFalse are the baseline for the model ,This means CLMSEX =28.95199 is the average number of Clmsex with no cy and no sy is 28.95199

d)

> 29.33265-2.81182

[1] 26.52083

This means MARTITAL =26.52083 is the average number of MARTITAL with no cy and no sy is 28.9519

CLMSEX\*CLMAGE

> model2=lm(LOSS~cy+sy+ATTORNEY+CLMSEX+factor(MARITAL)+CLMAGE+CLMSEX\*CLMAGE)

> summary(model2)

Call:

lm(formula = LOSS ~ cy + sy + ATTORNEY + CLMSEX + factor(MARITAL) +

CLMAGE + CLMSEX \* CLMAGE)

Residuals:

Min 1Q Median 3Q Max

-20.270 -4.550 -1.799 0.569 263.778

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 31.80300 5.11747 6.215 7.33e-10 \*\*\*

cyTRUE -2.02749 1.62507 -1.248 0.212

syTRUE -14.67013 3.55552 -4.126 3.97e-05 \*\*\*

ATTORNEY -6.39779 0.96229 -6.649 4.69e-11 \*\*\*

CLMSEX -1.95728 2.07474 -0.943 0.346

factor(MARITAL)2 -1.63862 1.14970 -1.425 0.154

factor(MARITAL)3 -3.56132 4.80602 -0.741 0.459

factor(MARITAL)4 0.60335 3.00373 0.201 0.841

CLMAGE -0.03204 0.09216 -0.348 0.728

CLMSEX:CLMAGE 0.04868 0.05664 0.860 0.390

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.7 on 1081 degrees of freedom

(249 observations deleted due to missingness)

Multiple R-squared: 0.06521, Adjusted R-squared: 0.05742

F-statistic: 8.378 on 9 and 1081 DF, p-value: 3.701e-12

E) 1.e1)

Ho: Beta9=0;

H1: Beta9!=0;

> dim(auto)

[1] 1340 8

> #1340-249=1091

> qt(.975,1081)

[1] 1.962161

> #critical t=1.962161

#Since absolute tscore = .860 <critical t,and also .39<alpha=.05

#we reject the null and believe this term is insignificant

f)

newdata=data.frame(ATTORNEY=1,CLMSEX=1,MARITAL=2, cy=TRUE,sy=FALSE,CLMAGE=35)

> predict(model2, newdata, interval="confidence", level=.95)

fit lwr upr

1 20.36448 12.68405 28.0449

if Xs has features described as above, we have 95% confidence that predicted LOSS would have an average of score at the range of (12.68405 28.0449) .