> cut=cut(op$Price,br=c(50,59,79,99))

> table(cut)

cut

(50,59] (59,79] (79,99]

12 12 10

> cut

[1] (50,59] (50,59] (50,59] (50,59] (50,59] (50,59] (50,59] (50,59] (50,59]

[10] (50,59] (50,59] (50,59] (59,79] (59,79] (59,79] (59,79] (59,79] (59,79]

[19] (59,79] (59,79] (59,79] (59,79] (59,79] (59,79] (79,99] (79,99] (79,99]

[28] (79,99] (79,99] (79,99] (79,99] (79,99] (79,99] (79,99]

Levels: (50,59] (59,79] (79,99]

> model1 = lm(Sales ~ factor(cut))

> summary(model1)

Call:

lm(formula = Sales ~ factor(cut))

Residuals:

Min 1Q Median 3Q Max

-2385.92 -585.29 37.97 657.58 1473.70

Coefficients:

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

(Intercept) 4151.7 253.6 16.371 < 2e-16 \*\*\*

factor(cut)(59,79] -1090.8 358.6 -3.042 0.00476 \*\*

factor(cut)(79,99] -2271.4 376.2 -6.039 1.1e-06 \*\*\*

---

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

Residual standard error: 878.5 on 31 degrees of freedom

Multiple R-squared: 0.5407, Adjusted R-squared: 0.5111

F-statistic: 18.25 on 2 and 31 DF, p-value: 5.78e-06

#Ho:Beta=0;

H1:Beta!=0;

#model1 is significant, because the F-pvalue <.05 on 31 degrees of freedom,

> 4151.7-1090.8

[1] 3060.9

B1=3060.9

2) when we increase price from 59 to 79,the sales is dropped by 1090.8 from 4151.7 at $59 as the bas line.b1=4151.7-10901.8

b.

PromotionL= (1\*( Promotion==200))

PromotionM= (1\*( Promotion==400))

PromotionH= (1\*( Promotion==600))

check = data.frame(Promotion, PromotionH, PromotionM, PromotionL)

fix(check)

> PromotionL= (1\*( Promotion==200))

> PromotionM= (1\*( Promotion==400))

> PromotionH= (1\*( Promotion==600))

>

> check = data.frame(Promotion, PromotionH, PromotionM, PromotionL)

> fix(check)

> model2=lm(Sales~PromotionH+PromotionM+PromotionL)

> summary(model2)

Call:

lm(formula = Sales ~ PromotionH + PromotionM + PromotionL)

Residuals:

Min 1Q Median 3Q Max

-1722.8 -804.5 -252.1 947.0 1743.2

Coefficients: (1 not defined because of singularities)

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

(Intercept) 2397.8 314.3 7.629 1.33e-08 \*\*\*

PromotionH 1666.5 466.2 3.575 0.00117 \*\*

PromotionM 597.0 444.5 1.343 0.18897

PromotionL NA NA NA NA

---

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

Residual standard error: 1089 on 31 degrees of freedom

Multiple R-squared: 0.2946, Adjusted R-squared: 0.2491

F-statistic: 6.474 on 2 and 31 DF, p-value: 0.004473

b.1) #Ho:Beta=0;

H1:Beta!=0;

#model1 is significant, because the F-pvalue .00447<.05 on 31 degrees of freedom,

> 2397.8+597

[1] 2994.8

> model4= lm(Sales ~ Promotion)

newdata = data.frame(Promotion= 400)

predict(model4, newdata, interval="confidence", level=.95)

fit lwr upr

1 3147.242 2769.697 3524.787

If an individual has 400 as its value in Promotion, we have 95% confidence that the predicted Y for that individual would have a value at (2769.697 3524.787) .

I would choose model1,because it has higher Rsqr and Rsqr adj,while F-p value is very small.

Thus,It would let X explain more of Y variable

#Price as the baseline for the model is intercept y

2.

a)