Yanjie Hu 1006750044 STA314 HW1

Problem 4

#(a)

```
library(ISLR)
attach(Carseats)
```

```
lm.fit=lm(Sales~Price+Urban+US)
```

#(b)

```
lm.fit
```

```
##
## Call:
## lm(formula = Sales ~ Price + Urban + US)
##
## Coefficients:
## (Intercept)        Price      UrbanYes         USYes
##    13.04347     -0.05446      -0.02192       1.20057
```

In this case, the intercept here is when the price is equals to 0, and not a US and urban car here, the mean sale is 13.04347. The price coefficient here is saying that when price is increased by 1 unit, the mean sale will decrease 0.05446 unit. When it is urban(when the price and us variable is fixed), then the mean sale will decrease by 0.02192 unit and when it is from US(when the price and urban is fixed) it will increase 1.20057 unit of sales.

#(c)

$$x_i1 = \begin{cases} 1 & \text{if it is from urban} \\ 0 & \text{if it is not from urban} \end{cases}$$

$$x_i2 = \begin{cases} 1 & \text{if it is from US} \\ 0 & \text{if it is not from US} \end{cases}$$

$$Y_i|x = \beta_0 + \beta_1 x_i + \varepsilon_i$$

#(d)

Price and US.

```
summary(lm.fit)
```

```
##
## Call:
## lm(formula = Sales ~ Price + Urban + US)
##
## Residuals:
##     Min     1Q  Median     3Q    Max
## -6.9206 -1.6220 -0.0564  1.5786  7.0581
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13.043469   0.651012  20.036  < 2e-16 ***
## Price       -0.054459   0.005242 -10.389  < 2e-16 ***
## UrbanYes    -0.021916   0.271650  -0.081    0.936
## USYes        1.200573   0.259042   4.635 4.86e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.472 on 396 degrees of freedom
## Multiple R-squared:  0.2393, Adjusted R-squared:  0.2335
## F-statistic: 41.52 on 3 and 396 DF,  p-value: < 2.2e-16
```

Since the p-value of price and USYes is smaller than 0.05, then they are signifiant, and we can reject the null hypothesis.

#(e)

```
lm.fit.new=lm(Sales~Price+US)
```

#(f)

```
summary(lm.fit)$r.squared
```

```
## [1] 0.2392754
```

```
summary(lm.fit.new)$r.squared
```

```
## [1] 0.2392629
```

Since we can see that the r squared is not shown big difference and the r square is not better when larger, because it will also letting larger whatever you put in any data sample.

#(g)

```
confint(lm.fit.new)
```

```
##                     2.5 %       97.5 %
## (Intercept) 11.79032020 14.27126531
## Price        -0.06475984 -0.04419543
## USYes         0.69151957  1.70776632
```

#(h)