Self-Assessment #3 - Solutions

MGT 6203 - Data Analytics in Business

Georgia Tech, Fall 2020

Install and load the dataset named Carseats (in the ISLR package) into R.

```
library("ISLR")
data("Carseats")

Model1 = lm(Carseats$Sales~Carseats$Price)
summary(Model1)

##
## Call:
## lm(formula = Carseats$Sales ~ Carseats$Price)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.5224 -1.8442 -0.1459  1.6503  7.5108
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    13.641915   0.632812  21.558   <2e-16 ***
## Carseats$Price -0.053073   0.005354  -9.912   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.532 on 398 degrees of freedom
## Multiple R-squared:  0.198,  Adjusted R-squared:  0.196
## F-statistic: 98.25 on 1 and 398 DF,  p-value: < 2.2e-16

print("Question1")

## [1] "Question1"

summary(Model1)$adj.r.squared

## [1] 0.195966

print("Question2")

## [1] "Question2"

summary(Model1)$coefficients[2]

## [1] -0.05307302

print("Question3")
```

```
## [1] "Question3"

summary(Model1)$coefficients[6]

## [1] -9.911997

print("Question4")

## [1] "Question4"

print("Since the p value is less that 0.05, the estimated Coefficient is stat
istically different from 0")

## [1] "Since the p value is less that 0.05, the estimated Coefficient is sta
tistically different from 0"

Carseats$Bad_Shelf<-ifelse(Carseats$ShelveLoc=="Bad",1,0)
Carseats$Good_Shelf<-ifelse(Carseats$ShelveLoc=="Good",1,0)

Model2<-lm(Carseats$Sales~Carseats$Price+Carseats$Bad_Shelf+Carseats$Good_She
lf)
summary(Model2)

##
## Call:
## lm(formula = Carseats$Sales ~ Carseats$Price + Carseats$Bad_Shelf +
##     Carseats$Good_Shelf)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8229 -1.3930 -0.0179  1.3868  5.0780
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         13.863824   0.487021  28.467  < 2e-16 ***
## Carseats$Price      -0.056698   0.004059 -13.967  < 2e-16 ***
## Carseats$Bad_Shelf  -1.862022   0.234748  -7.932 2.23e-14 ***
## Carseats$Good_Shelf  3.033825   0.245178  12.374  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.917 on 396 degrees of freedom
## Multiple R-squared:  0.5426, Adjusted R-squared:  0.5391
## F-statistic: 156.6 on 3 and 396 DF,  p-value: < 2.2e-16

print("Question5")

## [1] "Question5"

summary(Model2)$coefficients[3]

## [1] -1.862022
```

```
print("Question6")

## [1] "Question6"

summary(Model2)$coefficients[1]

## [1] 13.86382

print("Question7")

## [1] "Question7"

summary(Model2)$coefficients[1]+summary(Model2)$coefficients[3]

## [1] 12.0018

print("Question8")

## [1] "Question8"

print("The coefficient of Good_Shelf captures the difference in sales of cars
eats if they are located in the Good shelf location compared to that of carse
ats located in the Medium shelf location.")

## [1] "The coefficient of Good_Shelf captures the difference in sales of car
seats if they are located in the Good shelf location compared to that of cars
eats located in the Medium shelf location."

#setwd("~/Desktop/PriceDemand.csv")
PriceDemand = read.csv("PriceDemand.csv", header = TRUE)
Model3<-lm(PriceDemand$Qty~PriceDemand$Price)
summary(Model3)

##
## Call:
## lm(formula = PriceDemand$Qty ~ PriceDemand$Price)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -338.04 -153.96   -5.62  156.90  676.23
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3501.99     225.57  15.525  < 2e-16 ***
## PriceDemand$Price  -393.63      44.14  -8.918 9.38e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 214.3 on 48 degrees of freedom
## Multiple R-squared:  0.6236, Adjusted R-squared:  0.6158
## F-statistic: 79.52 on 1 and 48 DF,  p-value: 9.377e-12

print("Question9")
```

```
## [1] "Question9"

print("One dollar increase in price decreases demand by 394 units")

## [1] "One dollar increase in price decreases demand by 394 units"

PriceDemand$Price_ln<-log(PriceDemand$Price)
Model4<-lm(PriceDemand$Qty~PriceDemand$Price_ln)
summary(Model4)

##
## Call:
## lm(formula = PriceDemand$Qty ~ PriceDemand$Price_ln)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -319.34 -135.82   -0.56  147.41  560.01
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            4723.8      322.4   14.65  < 2e-16 ***
## PriceDemand$Price_ln  -1993.9      199.2  -10.01 2.46e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 198.8 on 48 degrees of freedom
## Multiple R-squared:  0.6761, Adjusted R-squared:  0.6693
## F-statistic: 100.2 on 1 and 48 DF,  p-value: 2.457e-13

#For small p, approximately log([100 + p]/100) ≈ p/100. For p = 1, this means
that b1/100 can be interpreted approximately as the expected increase in Y fr
om a 1% increase in X

print("Question10")

## [1] "Question10"

print("When price increases by 1%, quantity decreases by 19.94 units")

## [1] "When price increases by 1%, quantity decreases by 19.94 units"

PriceDemand$Qty_ln<-log(PriceDemand$Qty)
Model5<-lm(PriceDemand$Qty_ln~PriceDemand$Price)
summary(Model5)

##
## Call:
## lm(formula = PriceDemand$Qty_ln ~ PriceDemand$Price)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.245148 -0.091984 -0.000218  0.104754  0.264779
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.48937    0.13359  63.550  < 2e-16 ***
## PriceDemand$Price -0.23550    0.02614  -9.009 6.88e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1269 on 48 degrees of freedom
## Multiple R-squared:  0.6284, Adjusted R-squared:  0.6206
## F-statistic: 81.16 on 1 and 48 DF,  p-value: 6.877e-12
```

```
#Interpretation of the estimated coefficient b1 is that a one-unit increase i
n X will produce an expected increase in log Y of b1 units. In terms of Y its
elf, this means that the expected value of Y is multiplied by e^b1.
#For small values of b1, approximately e^b1 ≈ 1+b1. We can use this for the f
ollowing approximation for a quick interpretation of the coefficients: 100*b1
is the expected percentage change in Y for a unit increase in X.
print("Question11")
```

```
## [1] "Question11"
```

```
print("When price increases by $0.1, quantity decreases (on average) by 2.35%
")
```

```
## [1] "When price increases by $0.1, quantity decreases (on average) by 2.35
%"
```

```
Model6<-lm(PriceDemand$Qty_ln~PriceDemand$Price_ln)
summary(Model6)
```

```
##
## Call:
## lm(formula = PriceDemand$Qty_ln ~ PriceDemand$Price_ln)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23437 -0.08879 -0.00340  0.09432  0.20484
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            9.2011     0.1946  47.273  < 2e-16 ***
## PriceDemand$Price_ln  -1.1810     0.1202  -9.822 4.55e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.12 on 48 degrees of freedom
## Multiple R-squared:  0.6677, Adjusted R-squared:  0.6608
## F-statistic: 96.46 on 1 and 48 DF,  p-value: 4.552e-13
```

```r
#Increasing log(Price) by 0.01 changes log(Qty) by b1 * 0.01 units which impl
ies increasing Price by 1% changes Qty by b1 %
print("Question12")
```

```
## [1] "Question12"
```

```r
print("When price increases by 1%, quantity decreases (on average) by 1.18%")
```

```
## [1] "When price increases by 1%, quantity decreases (on average) by 1.18%"
```