SafeBabies Company

Mohamed Megahed

February 28, 2020

SafeBabies Problem

SafeBabies is a large company who is producing car seats for babies and toddlers. They sell their products all over the US and abroad. The management team has hired you as a Business Analytics consultant to help them maximizing their profit.

The primary task is to determine:

1. the optimal price for selling the car seats at those stores where the shelve location is good (i.e. the product is highly visible)?
2. the optimal price for selling the car seats at those stores where the shelve location is bad (i.e. the product is highly visible)?

The cost of producing each car seat is $55.0

3. Plot the optimal price for selling the car seats at those stores where the shelve location is good and those where the shelve location is bad when varying the production costs from $40 to $85.

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.6.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ISLR)
```

```
## Warning: package 'ISLR' was built under R version 3.6.2
```

Our Problem deals with only 3 variables "Sales, Price and ShelveLoc"

```
Safe_Baby <- Carseats[,c(1,6,7)]
head(Safe_Baby)
```

```
##   Sales Price ShelveLoc
## 1  9.50   120       Bad
```

```
## 2 11.22    83      Good
## 3 10.06    80    Medium
## 4  7.40    97    Medium
## 5  4.15   128       Bad
## 6 10.81    72       Bad
```

Understanding our dataset

```
str(Safe_Baby)
```

```
## 'data.frame':    400 obs. of  3 variables:
##  $ Sales    : num  9.5 11.22 10.06 7.4 4.15 ...
##  $ Price    : num  120 83 80 97 128 72 108 120 124 124 ...
##  $ ShelveLoc: Factor w/ 3 levels "Bad","Good","Medium": 1 2 3 3 1 1 3 2 3
3 ...
```

```
summary(Safe_Baby)
```

```
##      Sales            Price        ShelveLoc
##  Min.   : 0.000   Min.   : 24.0   Bad   : 96
##  1st Qu.: 5.390   1st Qu.:100.0   Good  : 85
##  Median : 7.490   Median :117.0   Medium:219
##  Mean   : 7.496   Mean   :115.8
##  3rd Qu.: 9.320   3rd Qu.:131.0
##  Max.   :16.270   Max.   :191.0
```

Where

- Sales is unit sales (in thousands) at each location
- Price is the price that company charges for car seats at each site, and
- ShelveLoc is a factor with levels Bad, Good and Medium indicating the quality of the shelving location for the car seats at each store.

We will multiply Sales variable by 1000 as it is in thousands .

```
S_Baby <- Safe_Baby %>%
  mutate(Sales = Sales * 1000)
summary(S_Baby)
```

```
##      Sales            Price        ShelveLoc
##  Min.   :    0   Min.   : 24.0   Bad   : 96
##  1st Qu.: 5390   1st Qu.:100.0   Good  : 85
##  Median : 7490   Median :117.0   Medium:219
##  Mean   : 7496   Mean   :115.8
##  3rd Qu.: 9320   3rd Qu.:131.0
##  Max.   :16270   Max.   :191.0
```

*** Since the Question asks about the optimal price where the shelve location is good and where it is bad, we will split the data to 2 datasets one for `Good shelve location` and the other for `Bad shelve location`

```r
Good_shelve <- S_Baby %>%
  filter(ShelveLoc == "Good")

Bad_shelve <- S_Baby %>%
  filter(ShelveLoc == "Bad")
```

Problem Formultion

The Optimal Price is the price Per Unit that can maximize the total profit Maximizing Total Profit means "Maximizing Sales and Price while Minimizing total production Cost"

$$TotalProfit = Sales * ProfitMargin$$

We assume a constant production cost per unit $55

Sales has a negative relation with price as increase in price causes a decrease in sales

So our target now is to find the highest sales with highest unit price

The (sales ~ price) function can be presented by a linear relation as follows:

$$Sales(y) = b_0 + b_1 * Price(x)$$

$$ProfitMargin = PricePerUnit - ProductionCostPerUnit(C)$$

So, we will substitue this in the Total Profit equation

$$TotalProfit(TP) = (b_0 + b_1 * Price) * (Price - C)$$

$ TP = b_0 * Price - b_0 * C + b_1 * Price^2 - b_1 * Price * C $

$$TP = b_1 * Price^2 + (b_0 - b_1 C)Price - b_0 C$$

We will take the first derivative of the quadratic equation and set the derivative to zero.

$$\frac{dTP}{dPrice} = 2b_1 Price + b_0 - b_1 C$$

$ 2 b_1Price + b_0 - b_1C = 0 $

$ 2 b_1Price = -b_0 + b_1C $

$$Price = \frac{-b_0 + b_1 C}{2b_1}$$

Setting an Optimal Price function using the above formula:

```r
Optimal_Price <- function(cost, b_0, b_1) {
    return(((-b_0) + (b_1 * cost)) / (2 * b_1))
}
```

Finding the Optimal Price for Good Shelve Locations:

```r
library(magrittr)
```

```
## Warning: package 'magrittr' was built under R version 3.6.2
```

```
Good_coefficients <- lm(Sales ~ Price, data = Good_shelve) %>%
  use_series("coefficients")
Good_coefficients
```

```
## (Intercept)       Price
## 17968.86360   -65.78477
```

```
Optimal_Price_Good <-
Optimal_Price(55,Good_coefficients[[1]],Good_coefficients[[2]])
Optimal_Price_Good
```

```
## [1] 164.0731
```

So, The Optimal Price for Good shelve Location is $164.0731

Finding the Optimal Price for Bad Shelve Locations:

```
Bad_coefficients <- lm(Sales ~ Price, data = Bad_shelve) %>%
  use_series("coefficients")
Bad_coefficients
```

```
## (Intercept)       Price
## 11832.98389   -55.22028
```

```
Optimal_Price_Bad <-
Optimal_Price(55,Bad_coefficients[[1]],Bad_coefficients[[2]])
Optimal_Price_Bad
```

```
## [1] 134.6435
```

So, The Optimal Price for Bad shelve Location is $134.6435

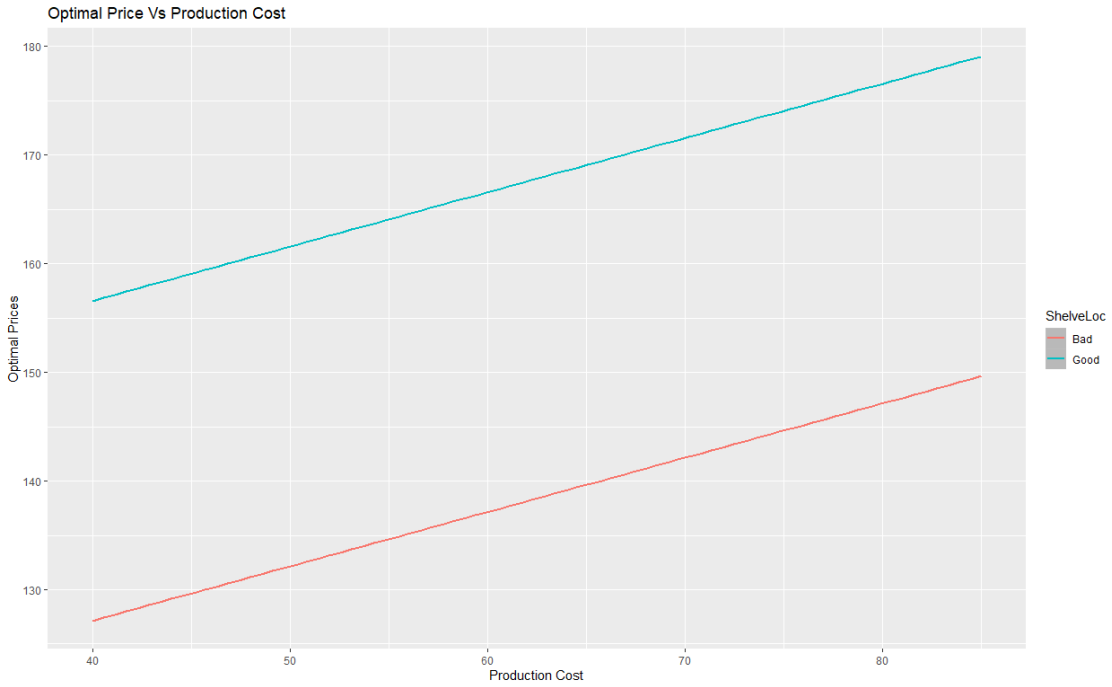
Plotting the optimal Price for Good and Bad Locations when varying the production costs from $40 to $85

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.6.2
```

```
prod_costs <- seq(40, 85, by = 5)
Good_Optimal_Prices1 <- c()
Bad_Optimal_Prices1 <- c()

    for (prod_cost in prod_costs)
      {
        Good_Optimal_Price <- Optimal_Price(prod_cost,
Good_coefficients[[1]], Good_coefficients[[2]])
        Good_Optimal_Prices1 <- c(Good_Optimal_Prices1, Good_Optimal_Price)

        Bad_Optimal_Price <- Optimal_Price(prod_cost, Bad_coefficients[[1]],
Bad_coefficients[[2]])
```

```
        Bad_Optimal_Prices1 <- c(Bad_Optimal_Prices1, Bad_Optimal_Price)
      }

    Optimal_Prices <- data.frame(
        Optimal_Good_Prices = Good_Optimal_Prices1,
        Optimal_Bad_Prices = Bad_Optimal_Prices1
    )

    ggplot(Optimal_Prices) +
        geom_smooth(aes(x = prod_costs, y = Optimal_Good_Prices, color =
"Good"), method = "loess") +
        geom_smooth(aes(x = prod_costs, y = Optimal_Bad_Prices, color =
"Bad"), method = "loess") +
        labs(
            color = "ShelveLoc",
            x = "Production Cost",
            y = "Optimal Prices",
            title = "Optimal Price Vs Production Cost"
        )
```



We can conclude from the above plot that:

1. The Optimal Prices for Good shelve location is higher than the Optimal Prices for Bad shelve location
2. There is a positive linear correlation between Production cost and Optimal Price