

Assignment 1

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 maximize their profit.

Your primary task is to determine:

- 1- The optimal price for selling the car seats at those stores where the shelf location is “good” (i.e., the product is highly visible)?
- 2- The optimal price for selling the car seats at those stores where the shelf location is “bad” (i.e., the product is poorly visible)?

You have been told that 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 shelf location is “good” and separately for those stores where the shelf location is “bad” when varying the production costs from \$40 to \$85 (in \$5 increments).

Data:

All that you have been given is the following dataset (SafeBabies dataframe) which contains three variables for 400 stores of SafeBabies.

Library(ISLR)

```
SafeBabies <- Carseats %>% select("Sales", "Price", "ShelveLoc")
```

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.

Deliverables:

A single PDF file which is the output of your notebook/markdown (or screenshots) that shows your code, output and plots, explanation, and logic, is sufficient.

Note: This assignment is similar to what you will get in reality. Of course, not in terms of the data (because you will get 300 additional irrelevant and noisy variables!), but more in terms of project expectations. You will be given a dataset and then management expectation! Remember, problem formulation sometimes is more complicated than solving the problem itself.

Hint 1: There may be several different ways to tackle the question. A simple approach that I recommend is to divide your data into three different sub-sets based on the ShelveLoc first before any modeling.

Hint 2: You should be fine using a simple ordinary least square (OLS) linear model (i.e., the `lm()` function in R). You can also use regularized linear models if you want, though (see hint 3)

Hint 3: If you needed to build a generalized linear model (such as lasso or ridge) with a single predictor, be aware of the following issue with glmnet (I cannot nag enough with the design of this package) and look at the solution (well, more like a hack than the solution)

<https://stackoverflow.com/questions/46698836/how-to-perform-lasso-regression-with-a-single-predictor-column>

Good luck