**Adarsh Challa**

**Yousif Adnan Al Ghafli**

**Rahil Richard Pereira**

# Team Assignment 2

## Question 1 (30 points)

GoodMorning is a new line of probiotic juice products. Since its first product launch in 2015, GoodMorning were now on the shelves of nationwide retailers. GoodMorning had the challenge of raising product awareness, particularly about the benefit of probiotics and the great taste. As a recent stat-up in the fairly new probiotic market, GoodMorning did not have the funding to place nationwide advertisements. It instead allocated much of its small marketing budget to in-store demonstrations.

During in-store demonstration, GoodMorning representatives handed our product samples. The representative arrived at a specified store, set up a table, distributed samples, informed consumers about the product, and offered coupons to inspire purchase.

Another promotional program involved competitions among the five GoodMorning sales representatives for the most endcap displays. The endcap is the hub at the end of an aisle-one of the store’s most popular locations. Sales representatives competed for the highest number of stores they could convince to place GoodMorning’s products at the endcap. The winning sales representative received a big screen television. There was also a competition for the best-decorated endcap. The winning store received cases of product for the employees and gift cards.

The company’s in-store demo was launched in November of 2016. Due to limited marketing resources, management was pressured to cut any marketing expense that did not directly contribute to GoodMorning’s results. By July of 2017, several concerns were raised within the company about the effectiveness of the in-store demo program. Some questioned whether the demos boosted sales at all, while others were concerned that any boost was only temporary and that sales would revert to normal levels shortly afterward. Some executives questioned whether the increase in sales volume could justify the associated costs.

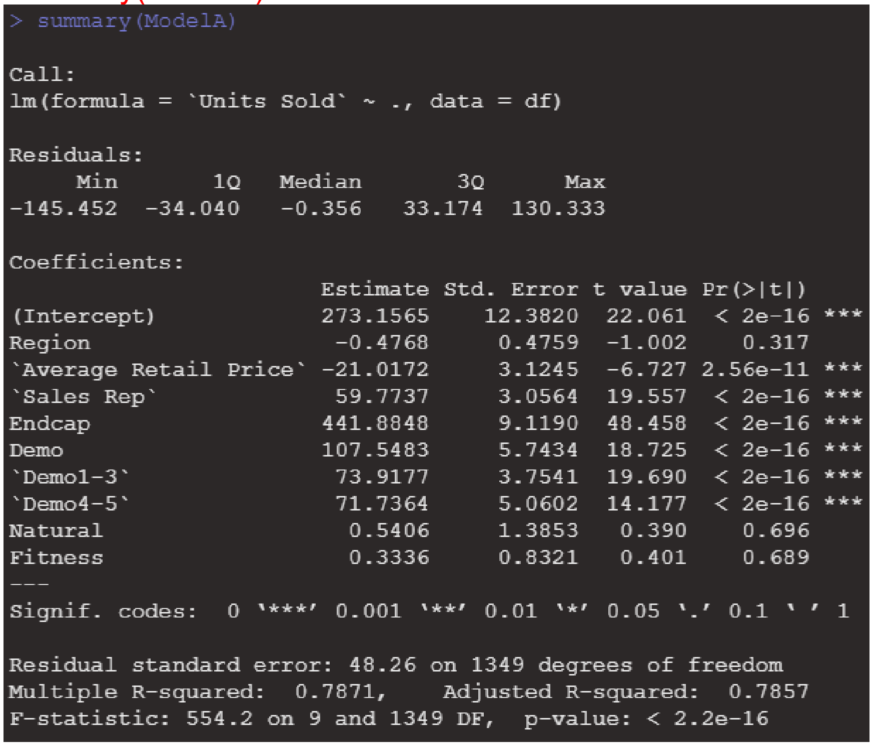
At the senior manager meeting, GoodMorning management asked Jim Martin, GoodMorning’s Marketing manager to justify the demo and endcap activities. Jim returned to his computer after the meeting and pored over the sales and promotion spreadsheet from the last few months. He recognized that statistics could be used to help his case. He decided to apply regression analysis to the sales and promotion data (GoodMorning.xlsx) to enable a decision on whether or not the company should continue its promotional programs.

### Model A:

Build a regression model with all variables in the data to explain the relationship between sales and promotional efforts. Let us refer to this model as Model *A*. Create the residual plot and the scatter plot of fit vs. UnitsSold.

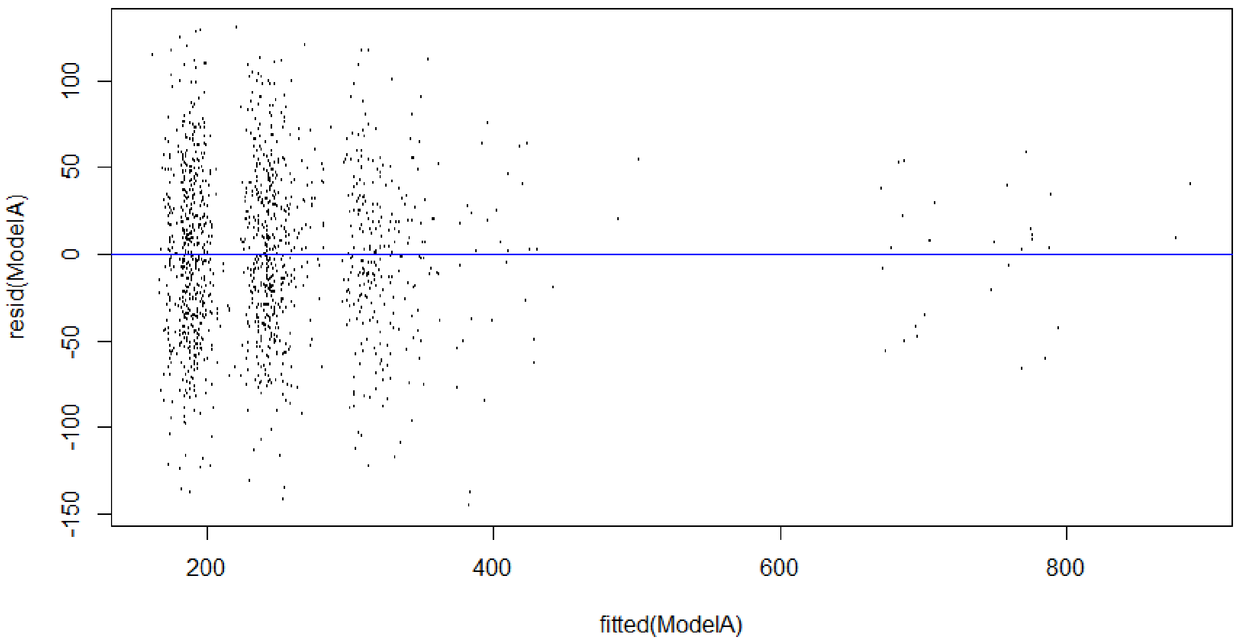
a) (10 points) Copy and paste the R code, the regression output, and the plots.

```
#Import dataset
library(readxl)
df <- read_excel("GoodMorning.xlsx")
View(df)
attach(df)

#MODEL A----

###Regression----
ModelA<-lm(`Units Sold`~.,data=df)
summary(ModelA)
```

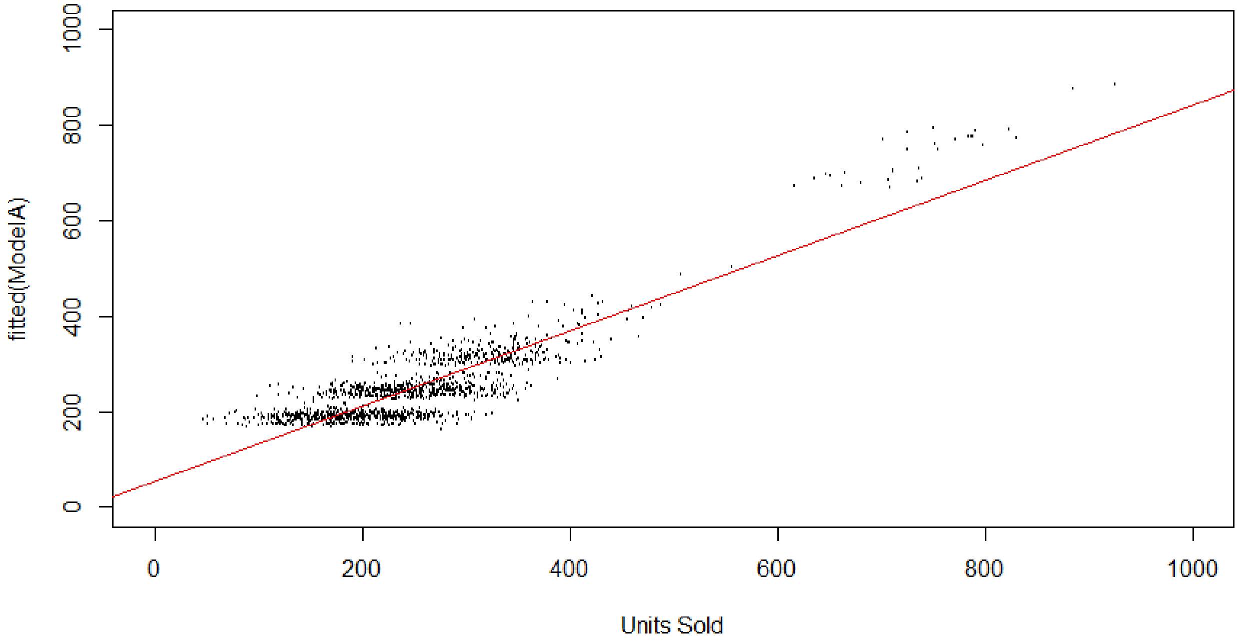


```
> summary(ModelA)

Call:
lm(formula = `Units Sold` ~ ., data = df)

Residuals:
     Min       1Q   Median       3Q      Max 
-145.452  -34.040   -0.356   33.174  130.333 

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)            273.1565    12.3820  22.061  < 2e-16 ***
Region                  -0.4768     0.4759  -1.002    0.317    
`Average Retail Price` -21.0172     3.1245  -6.727 2.56e-11 ***
`Sales Rep`             59.7737     3.0564  19.557  < 2e-16 ***
Endcap                 441.8848     9.1190  48.458  < 2e-16 ***
Demo                   107.5483     5.7434  18.725  < 2e-16 ***
`Demo1-3`               73.9177     3.7541  19.690  < 2e-16 ***
`Demo4-5`               71.7364     5.0602  14.177  < 2e-16 ***
Natural                  0.5406     1.3853   0.390    0.696    
Fitness                  0.3336     0.8321   0.401    0.689    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 48.26 on 1349 degrees of freedom
Multiple R-squared:  0.7871,	Adjusted R-squared:  0.7857 
F-statistic: 554.2 on 9 and 1349 DF,  p-value: < 2.2e-16
```

```
###Plot Residuals----
plot(resid(ModelA)~fitted(ModelA),cex=0.2)
abline(h=0, col="blue")
```



###Scatter Plot----
plot(`Units Sold`, fitted(ModelA), xlim = c(0,1000), ylim = c(0,1000), cex=0.2)
abline(lm(fitted(ModelA)~`Units Sold`), col="red")



b) (4 points) Discuss the performance and validity of the model, and how to improve and refine the model.

Adjusted R-Square:

Adjusted R-square of Model A is 78.57% which means that the model explains 78.57% of the variance in the dependent variable Given that this is an exploratory model with no tuning or transformation, this is fairly good model performance in terms of adjusted R-square.

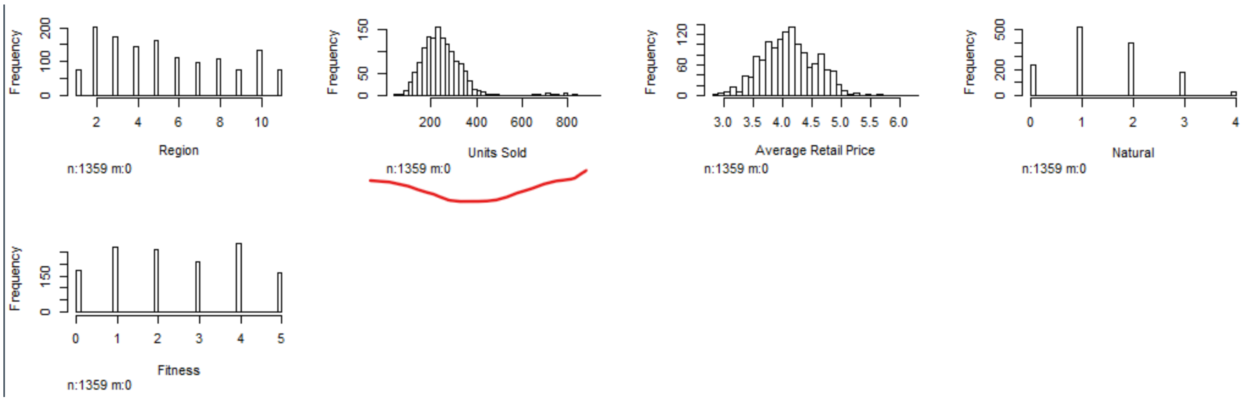
P-Value:

The p-value of the entire model is extremely small, showing the model is significant. For independent variables, the P-values of region, natural and fitness are high. These seem to be insignificant variables in the model and a better model could be built by removing them.

Distribution:

It can be observed that the dependent variable, Units Sold is extremely right-skewed. Taking the log of this variable in future models may be considered to see if model performance increases with normal distribution of this data.

```
###distribution----
install.packages("Hmisc")
library(Hmisc)
hist.data.frame(df)
```

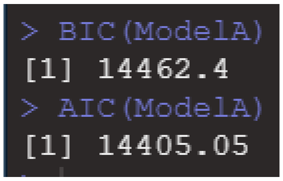


BIC and AIC:

It is recommended to use a stepwise model to see if these current  values can be reduced and help improve the model.

```
#BIC and AIC
BIC(ModelA)
```

AIC(ModelA)

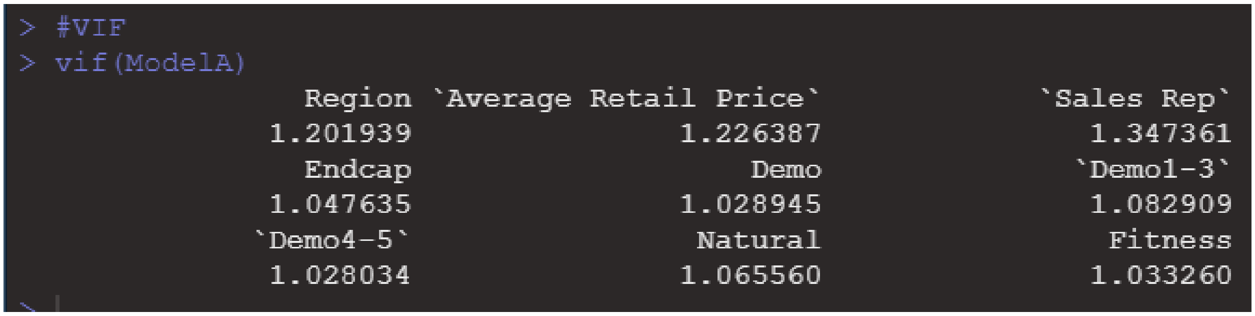


```
> BIC(ModelA)
[1] 14462.4
> AIC(ModelA)
[1] 14405.05
```

VIF:

All VIF values are around 1. This means that there is no multicollinearity between independent variables.

#VIF
vif(ModelA)



```
> #VIF
> vif(ModelA)
                 Region `Average Retail Price`            `Sales Rep`
               1.201939               1.226387               1.347361
                 Endcap                   Demo              `Demo1-3`
               1.047635               1.028945               1.082909
              `Demo4-5`                Natural                Fitness
               1.028034               1.065560               1.033260
```

Recommendations for improvement and refinement:

1. Try normalizing the distribution of the dependent variable by taking log.
2. Remove variables with high p-value since they are insignificant to model performance.
3. Try stepwise functions to lower the BIC and AIC values and for model selection
4. Create dummy variables of relevant categorical attributes and check if interaction terms are needed to improve model performance.
5. Review outliers in the dependent variable (seen in the scatter plot) and check if they can be segmented by end-cap.

These are all explored in Model B.

**Model B:**

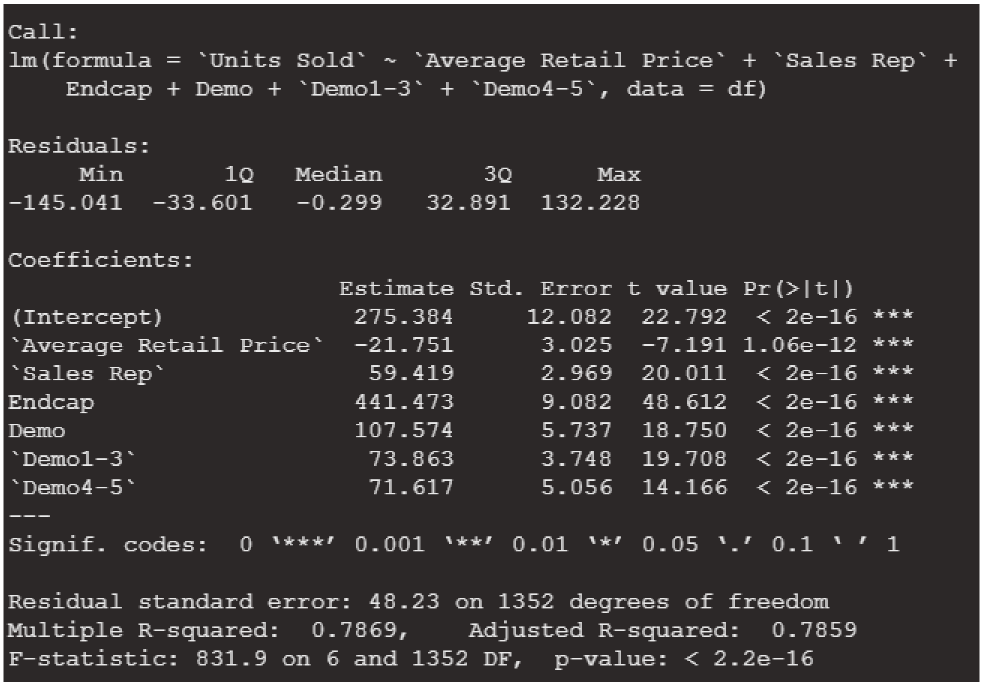
Build the best valid regression model to explain the relationship between sales and promotional efforts. You may use any transformation of your variables. Let us refer to this model as Model *B*. Create the residual plot and the scatter plot of fit vs. UnitsSold.

**STEPWISE:**

As per stepwise model selection, the attributes that should be kept in model B are:
Average retail price, sales rep, end cap, demo, demo 1-3, demo 4-5

```
#Stepwise for model selection
###Backward
stepBW <- step(ModelA, direction='backward', scope=formula(ModelA))
summary(stepBW)

##Forward
#define intercept-only model
intercept_only <- lm(`Units Sold` ~ 1 , data=df)
stepFW<- step(intercept_only, direction='forward', scope=formula(ModelA))
summary(stepFW)

##Both
stepBOTH<-step(ModelA, direction="both")
summary(stepBOTH)
```

The stepwise functions show the same optimal model selection:



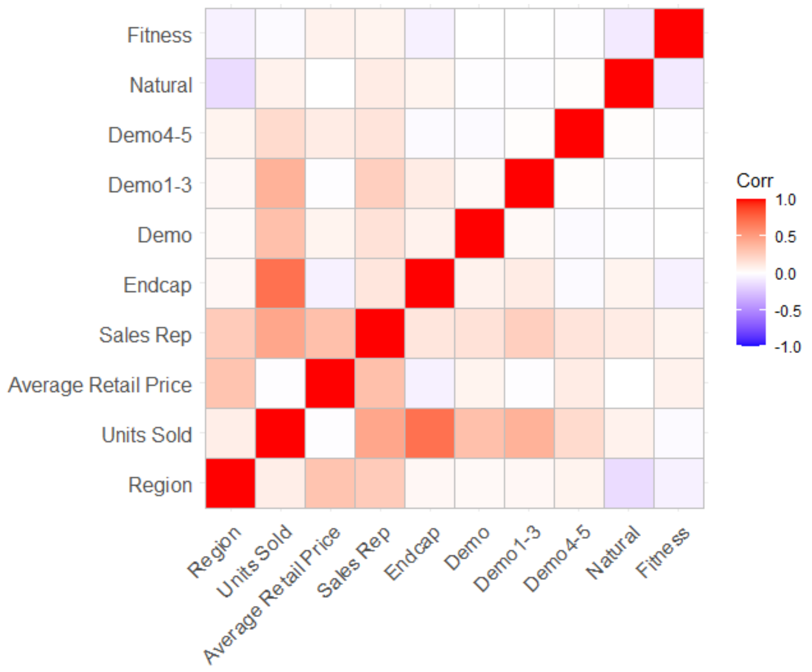
```
Call:
lm(formula = `Units Sold` ~ `Average Retail Price` + `Sales Rep` + 
    Endcap + Demo + `Demo1-3` + `Demo4-5`, data = df)

Residuals:
     Min       1Q   Median       3Q      Max 
-145.041  -33.601   -0.299   32.891  132.228 

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)             275.384     12.082  22.792  < 2e-16 ***
`Average Retail Price`  -21.751      3.025  -7.191 1.06e-12 ***
`Sales Rep`              59.419      2.969  20.011  < 2e-16 ***
Endcap                  441.473      9.082  48.612  < 2e-16 ***
Demo                    107.574      5.737  18.750  < 2e-16 ***
`Demo1-3`                73.863      3.748  19.708  < 2e-16 ***
`Demo4-5`                71.617      5.056  14.166  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 48.23 on 1352 degrees of freedom
Multiple R-squared:  0.7869,    Adjusted R-squared:  0.7859 
F-statistic: 831.9 on 6 and 1352 DF,  p-value: < 2.2e-16
```

**CORRELATION:**

Check correlation to see relationships between the variables.

```
###correlation----
install.packages("ggcorrplot")
```

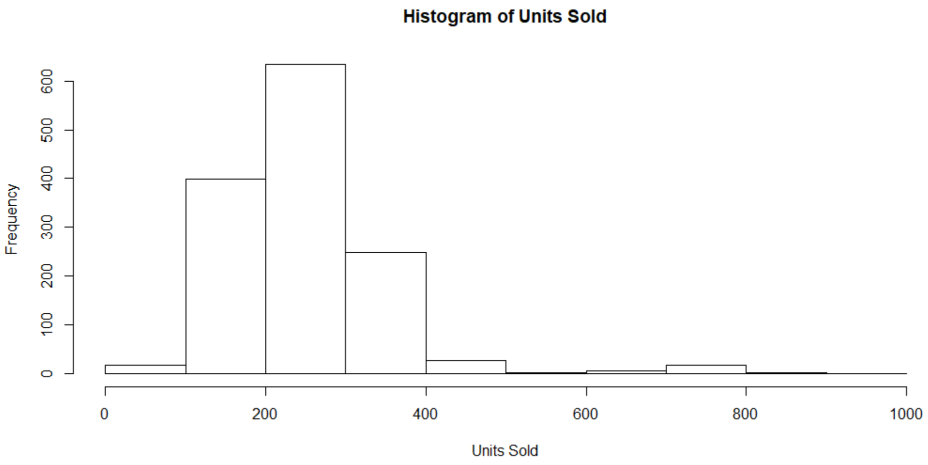
```
library(ggcorrplot)
ggcorrplot(cor(df))
```



## LOG:

Take log of units sold to normalize data distribution. Doing this reduced model performance significantly and was hence left out of the final model B.

```
##normalizing dist
log_units_sold<- log(`Units Sold`)
hist(log_units_sold)
hist(`Units Sold`)
```



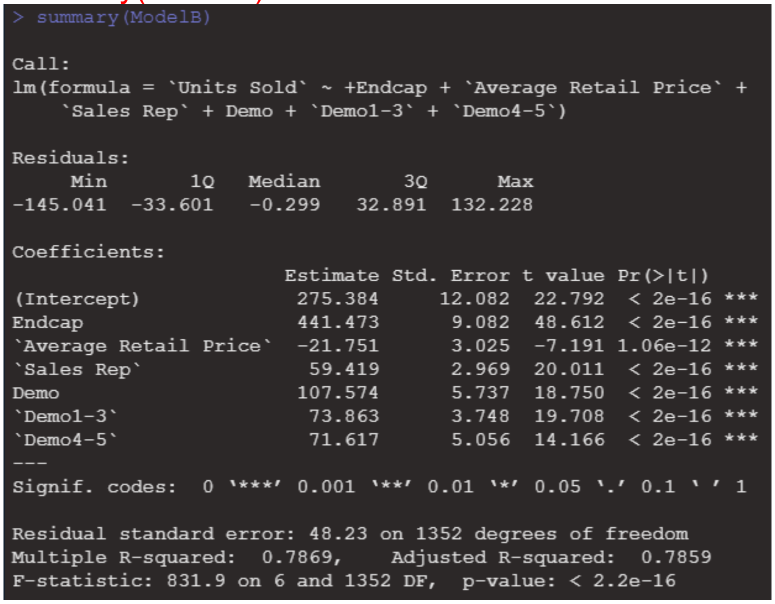
## Dummy variables and interaction terms:

These were considered while testing models and their performance. It was concluded that they did not add any significant improvement to R-square and other metrics used to evaluate model performance. In fact, it was observed that these seemed to deteriorate the model performance.

According to the principle of parsimony, we decided to explain the most with the least variables and utilize the domain knowledge acquired in the questions description.

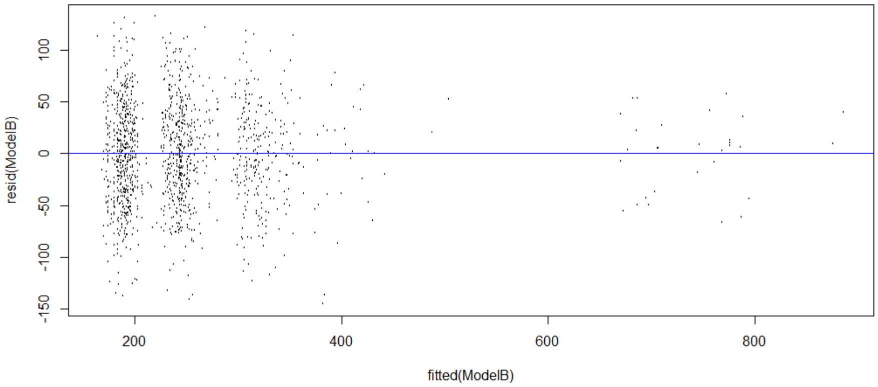
a) (6 points) Copy and paste the regression output, and the plots.

```
ModelB<-lm(formula= `Units Sold`~ +`Endcap`+`Average Retail Price`+`Sales
Rep`+Demo+`Demo1-3`+`Demo4-5`)
summary(ModelB)
```

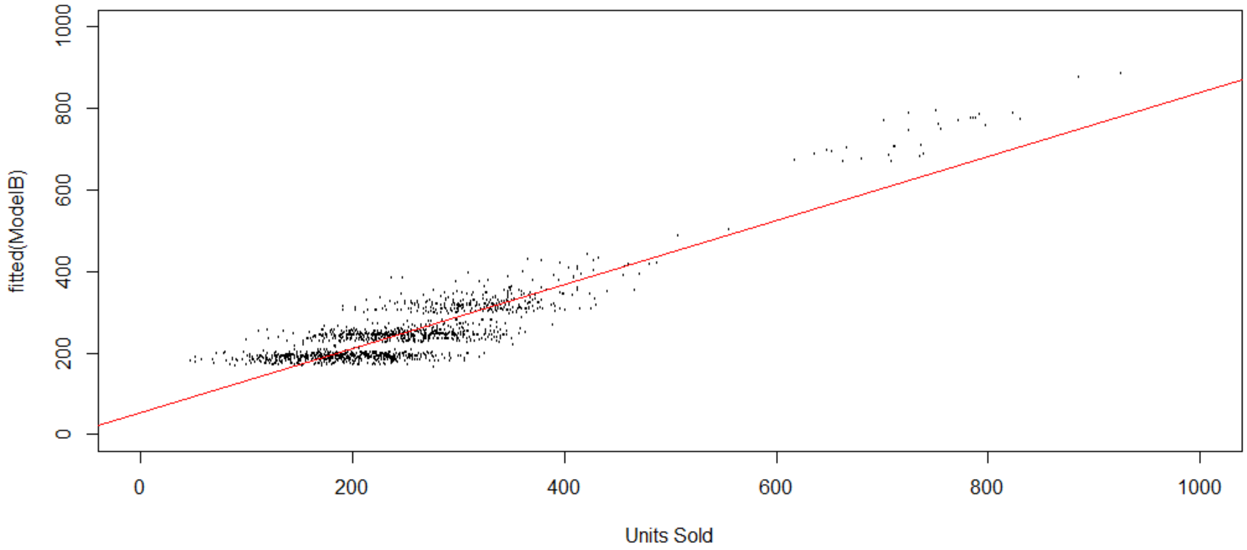


```
> summary(ModelB)

Call:
lm(formula = `Units Sold` ~ +Endcap + `Average Retail Price` +
    `Sales Rep` + Demo + `Demo1-3` + `Demo4-5`)

Residuals:
     Min       1Q   Median       3Q      Max
-145.041  -33.601   -0.299   32.891  132.228

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)             275.384     12.082  22.792  < 2e-16 ***
Endcap                  441.473      9.082  48.612  < 2e-16 ***
`Average Retail Price`  -21.751      3.025  -7.191 1.06e-12 ***
`Sales Rep`              59.419      2.969  20.011  < 2e-16 ***
Demo                    107.574      5.737  18.750  < 2e-16 ***
`Demo1-3`                73.863      3.748  19.708  < 2e-16 ***
`Demo4-5`                71.617      5.056  14.166  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 48.23 on 1352 degrees of freedom
Multiple R-squared:  0.7869,    Adjusted R-squared:  0.7859
F-statistic: 831.9 on 6 and 1352 DF,  p-value: < 2.2e-16
```

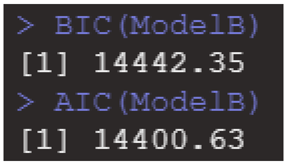
```
###Plot Residuals----
plot(resid(ModelB)~fitted(ModelB),cex=0.2)
abline(h=0, col="blue")
```



```
###Scatter Plot----
plot(`Units Sold`, fitted(ModelB), xlim = c(0,1000), ylim = c(0,1000), cex=0.2)
abline(lm(fitted(ModelB)~`Units Sold`), col="red")
```



```
#BIC and AIC

BIC(ModelB)

AIC(ModelB)
```



```
> BIC(ModelB)
[1] 14442.35
> AIC(ModelB)
[1] 14400.63
```

b) (2 points) Discuss the validity of the model.

Adjusted R-squared :

ModelA :  0.7857

ModelB :  0.7859

The adjusted R-squared value increased by 0.02 units after removing 3 independent variables. Usually , the R-squared value tends to decrease when variables are removed but the R-squared value increases by 0.02 units. This shows improvement in the model (when the independent variables Region , Natural and Fitness are removed). Having a higher R-square while using a lower number of independent variables to explain the proportion of variance in the dependent variable is a good measure of saying that Model B’s performance is better than Model A’s.

P-Value:

The p-values of all the dependent variables in the ModelB is extremely small, showing that the included variables give the information of the variance of the Dependent Variable, Units Sold. The variables that were removed in Model B, showed high p values in ModelA. Thus showed that they were insignificant to the model performance and hence were removed in Model B.

AIC and BIC:

AIC:

ModelA- 14405.05

ModelB  -14400.63

BIC:

ModelA - 14462.40

ModelB  - 14442.35

When we compare the ModelA and ModelB we see that there is a reduction in the both BIC and AIC values . General standard is that if there is reduction in the AIC and BIC values  there is improvement in the model. We can see that BIC reduced by 20 units and AIC reduced by 5 units.

Variance Inflation Factor :

```
> vif(ModelB)
        Endcap `Average Retail Price`       `Sales Rep`           Demo
       1.040136            1.150458          1.273026          1.027835
      `Demo1-3`           `Demo4-5`
       1.080418            1.027302
> vif(ModelA)
        Region `Average Retail Price`      `Sales Rep`          Endcap
       1.201939            1.226387          1.347361          1.047635
          Demo          `Demo1-3`         `Demo4-5`          Natural
```

```
1.028945        1.082909        1.028034        1.065560
 Fitness
1.033260
```

When we compare the VIF of ModelB with ModelA, the values for ModelB tend to be much closer compared to ModelA . This is a good sign of non-multicollinearity of the model.

Based on your model answer the following questions. Reference any tables/figures that you need to make your point:

c) (2 points) Does the in-store demo program boost the sales? If so, for how long does the sales lift last? Explain your answer.

Yes, the in-store demo boosts sales. Using the coefficients for Demo, Demo 1-3 and Demo 4-5, we can evaluate how impactful the  in-store demo program is.

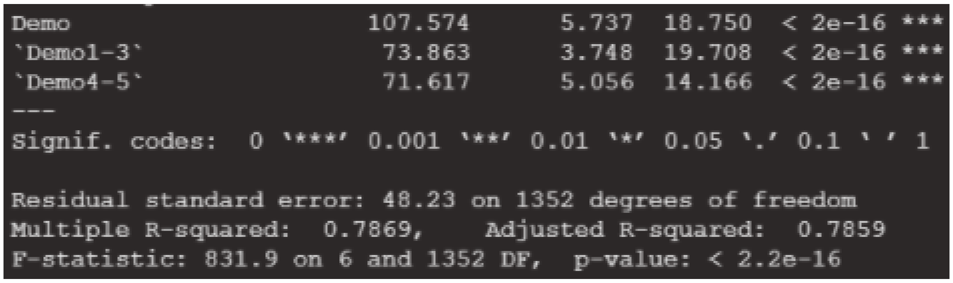
Demo: With all else held equal, when Demo happens in the current week, sales (units sold) are boosted by ~108 units.

Demo 1-3: With all else held equal, when Demo happens in the 1-3 weeks ago, sales (units sold) are boosted by ~74 units.

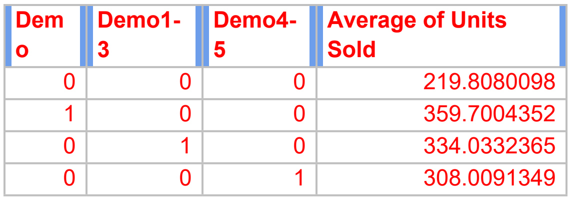
Demo 4-5: With all else held equal, when Demo happens in the 4-5 weeks ago, sales (units sold) are boosted by ~72 units.

It can be observed that the sales lift is highest when the in-store demo happens in the current week. Though the sales lift reduces in Week 1-3 (~74 units), there is still strong benefit to the sales than not having the in-store demo at all. Lastly, in the final weeks 4-5, the impact on sales reduces marginally to ~72 units. This decrease is very marginal.

In essence, the sales lift is greatest the current week and then there is a dip. But after the dip, the sales lift stabilizes in the final weeks showing there is continued impact towards the end of the in-store demo program. We recommend keeping the in-store demo program due to the high sales lift on units sold and continue reaping the benefits in subsequent weeks.



```
Demo                    107.574       5.737   18.750   < 2e-16 ***
`Demo1-3`                73.863       3.748   19.708   < 2e-16 ***
`Demo4-5`                71.617       5.056   14.166   < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 48.23 on 1352 degrees of freedom
Multiple R-squared:  0.7869,    Adjusted R-squared:  0.7859
F-statistic: 831.9 on 6 and 1352 DF,  p-value: < 2.2e-16
```



| Demo | Demo1-3 | Demo4-5 | Average of Units Sold |
|---|---|---|---|
| 0 | 0 | 0 | 219.8080098 |
| 1 | 0 | 0 | 359.7004352 |
| 0 | 1 | 0 | 334.0332365 |
| 0 | 0 | 1 | 308.0091349 |

Yes, in the final model , we can see that the p-values for the Demo variables( Demo, Demo1-3 and Demo4-5 ) are quite less.

d) (2 points) Does the placement of the product within the store (endcap promotion) affect the sales? Explain your answer.

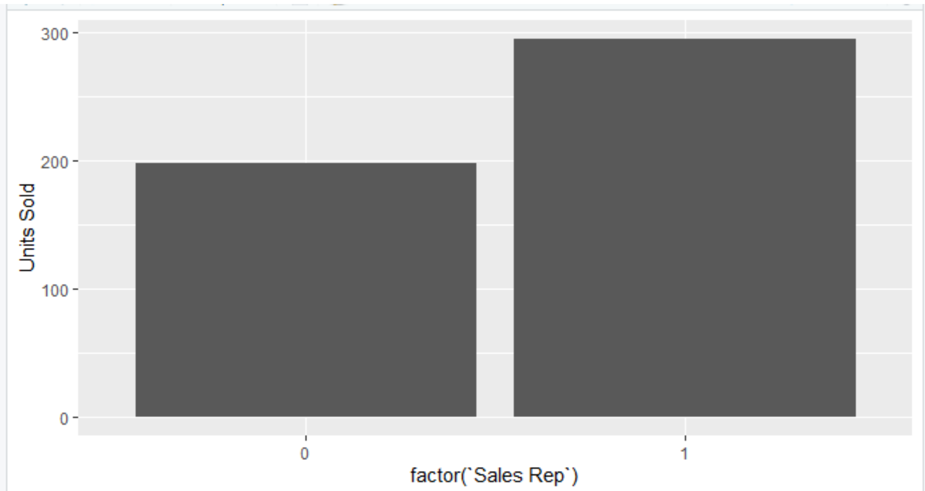
Yes, the endcap has the highest coefficient of all variables. This means that when the endcap promotion happens and all else held equal, the store in general will sell approximately 441 additional units. In addition, the t-score of the endcap is 48.612, highest between the independent variables, which indicates it has significant explanatory measure to the number of units sold.

e) (2 points) What other factors affect the sales of GoodMorning product? Explain your answer.

Sales Rep:

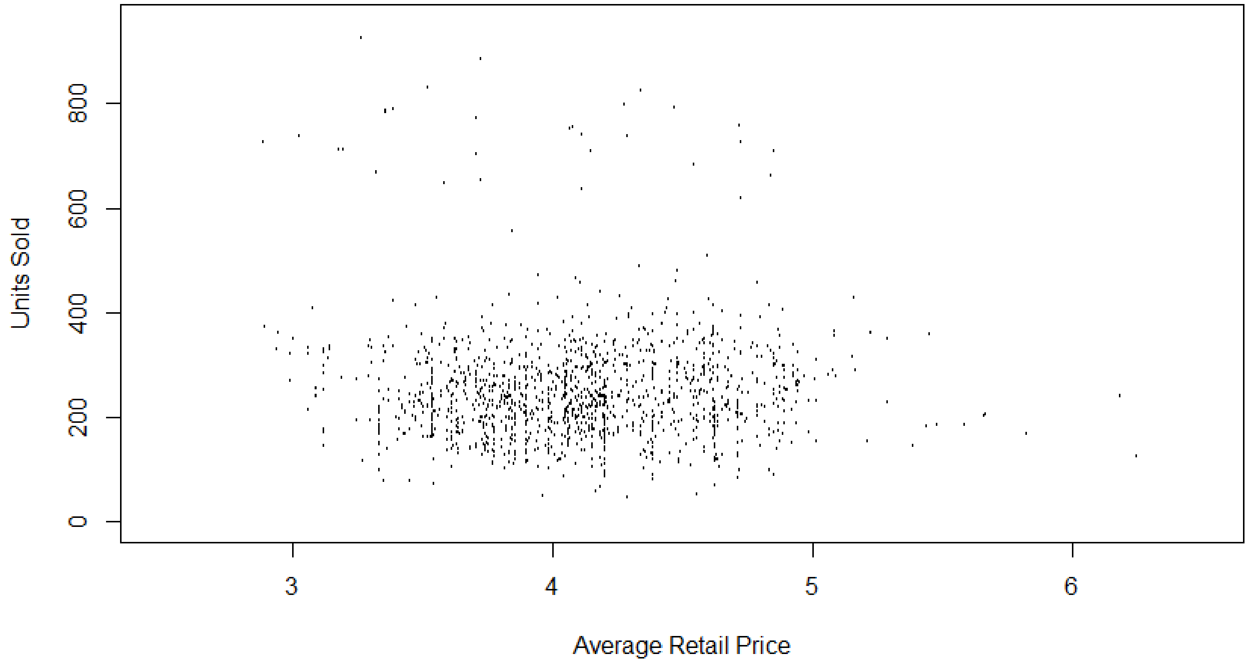
As shown in the bar plot below, having a sales rep shows to have significance on the units sold. Not having a sales rep (sales rep=0) shows fewer units sold.

```
#bar plot of sales rep and units sold
library(ggplot2)
ggplot(df, aes(x = factor(`Sales Rep`), y = `Units Sold`)) +  stat_summary(fun = "mean", geom = "bar")
```



Average Retail price:

```
#Scatter of average retail price and units sold.
plot(`Average Retail Price`,`Units Sold`, xlim = c(2.5,6.5), ylim = c(0,950), cex=0.2)
```



When we see the graph we can find that Units Sold has sometimes reached more than 600 units when the price is less than $5. When the price is more than $5 there are no traces where the No. of units sold has been more than 600 units. We can say that when the price is less than $5 the No. of Units sold is eventually more.

f) (2 points) Based on the regression output, what are your recommendations to GoodMorning management?

Demo program: We recommend keeping the in-store demo program due to the high sales lift on units sold. We can continue reaping these benefits in subsequent weeks as well.

Sales Rep: Having a sales rep at stores seems to have a significant impact on units sold so it is recommended to have sales reps in the stores.

Natural/Fitness: Since the proximity to these types of stores does not have a strong impact on sales, we can disregard these factors moving forward in the promotion of the product.

Region: There is low impact of the region to the number of units sold, hence, not included in the model. We do not recommend the management to focus their marketing effort on any specific region to lift the volume of units sold.

Average Sales Price: Seeing the trend in average price and units sold, it seems like sales are consistent when the average price is between $3.5 - $5. We can utilize this for profit maximization and pricing strategy.

Endcap: There is a significant relationship between endcap and units sold. Utilizing the endcap strategy is key to marketing the product and should be continued.

**Question 2 (30 points)** Use “Q2 data.xlsx” file.

A company is considering whether to market a new product. Assume, for simplicity, that if this product is marketed, there are only two possible outcomes: success or failure. The company assesses that the probabilities of these two outcomes are p and (1-p) respectively. If the product is marketed and it proves to be a failure, the company will have a net loss of $450,000. If the product is marketed and it proves to be a success, the company will have a net gain of $750,000. If the company decides not to market the product, there is no gain or loss.

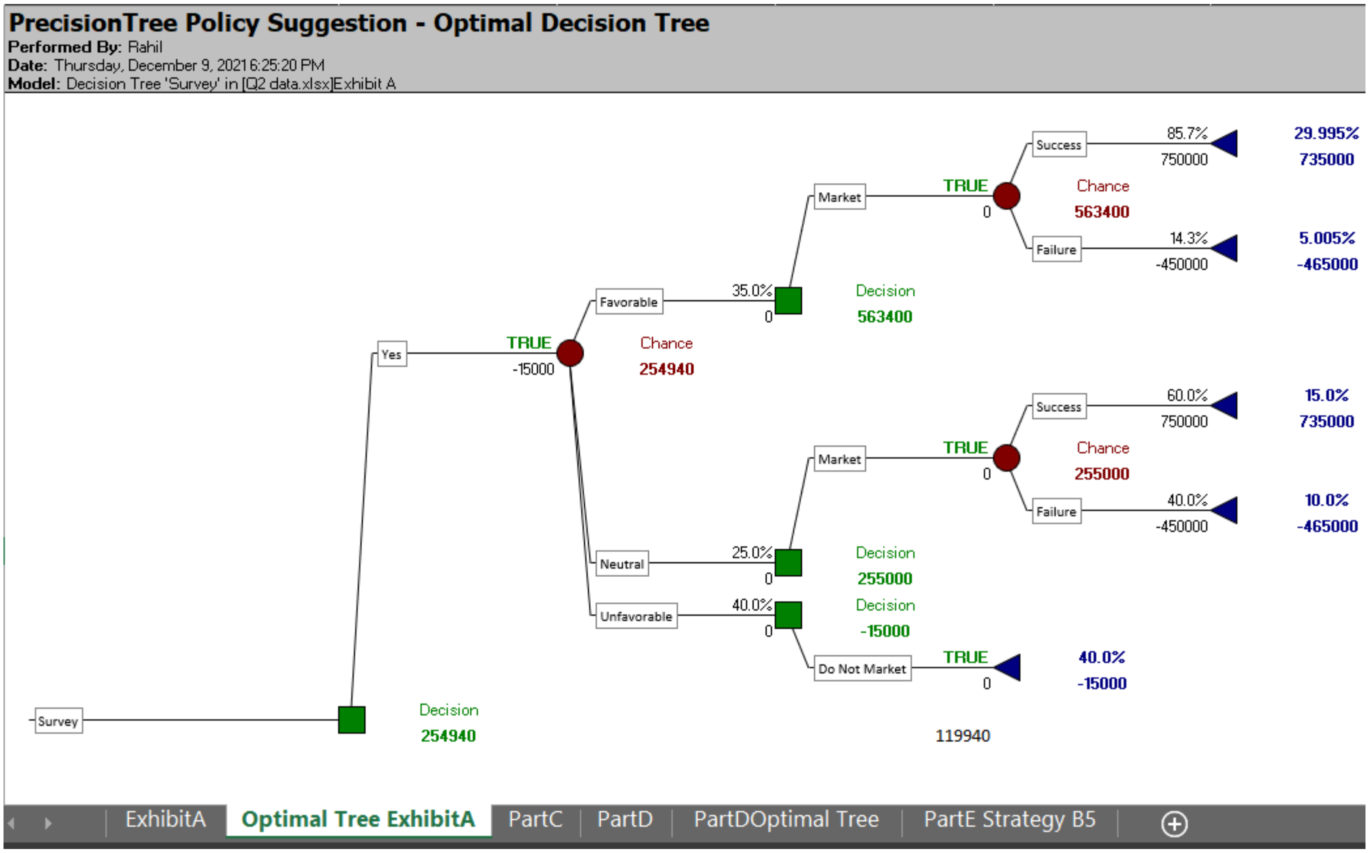
The company can first survey prospective buyers of this new product. The results of the consumer survey can be classified as favorable, neutral, or unfavorable. Based on similar survey for previous products, the company assesses these probabilities of favorable, neutral, or unfavorable results to be 0.6, 0.3, and 0.1 for product that will eventually be a success, and it assesses these probabilities to be 0.1, 0.2, and 0.7 for a product that will eventually be a failure. The total cost of administering this survey is C dollars.

Let p=0.5 and C= $15,000.

The company wants to construct a decision tree for this problem. The first step is to compute the posterior probabilities that the product will be eventually success and failure using the result from the consumer survey. The probabilities are given in Exhibit A.

The company would like to find the strategy that maximizes the company’s expected net earnings (EMV).

a) (10 pts) Construct a decision tree for this problem (Exhibit A). Generate the optimal decision strategy tree and paste the copy on your word document.



**PrecisionTree Policy Suggestion - Optimal Decision Tree**
**Performed By:** Rahil
**Date:** Thursday, December 9, 2021 6:25:20 PM
**Model:** Decision Tree 'Survey' in [Q2 data.xlsx]Exhibit A

- Does the optimal strategy involve conducting the survey?

Yes , The optimal strategy involves conducting the survey. The optimal tree shows this.

- What is the EMV under the optimal strategy?

The EMV under the optimal conditions is $254,940. The decision tree shows this and it can also be calculated by adding the probabilities and their corresponding values in the optimal decision tree:

735000*85.7% + 5.005%*(-465000) + 15%*735000 + 10%*-465000 +40%*(-465000)= 254940

b) (5 pts) Suppose that the total cost of administering this survey is $50,000.

- Does this change the company's decision?

Here , Expected Value of Information(EVI ) = EMV with (free) information - EMV without information

=> EVI = (255000+15000)-150000

= $120000

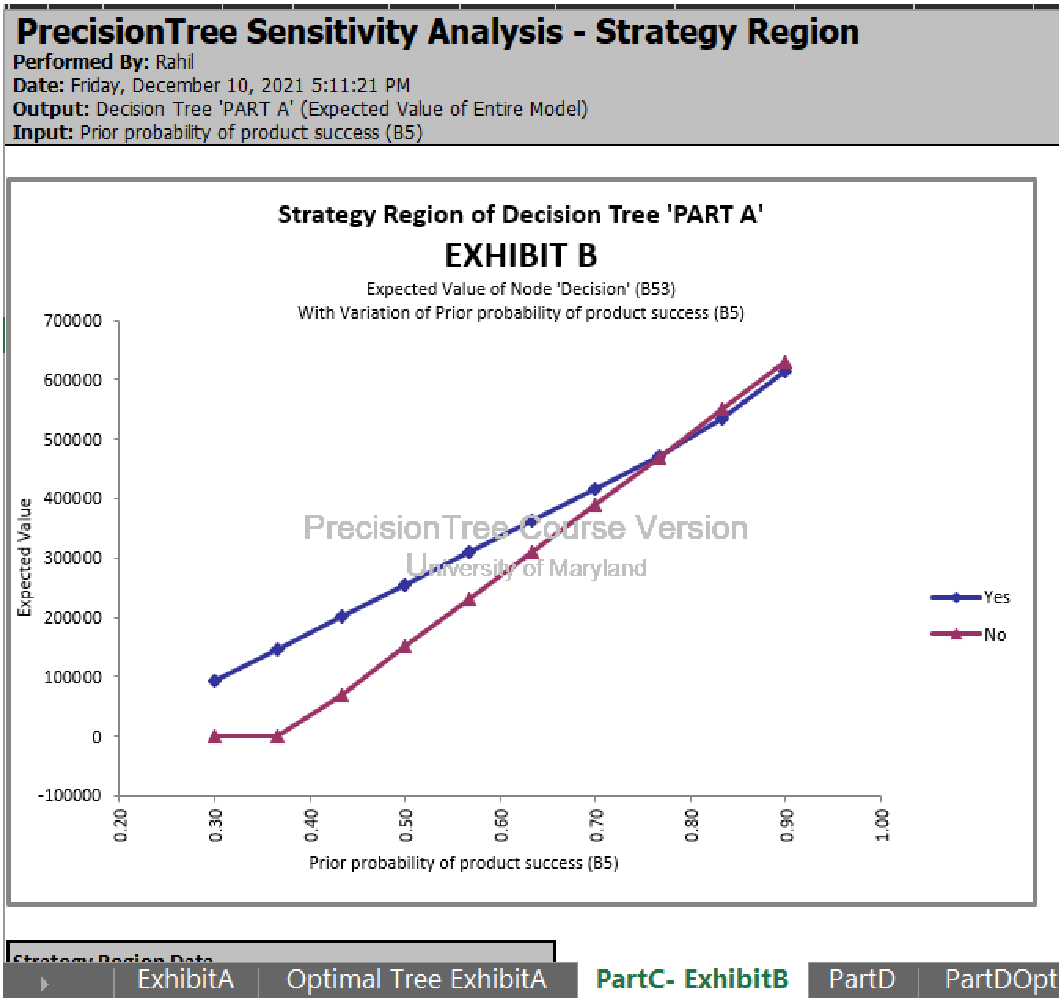
where EMV is Expected Monetary Value

Since, $50,000 is less than the EVI i.e. $120,000 the optimal decision doesn't change.

- What is the maximum amount that the company is willing to pay for the survey?

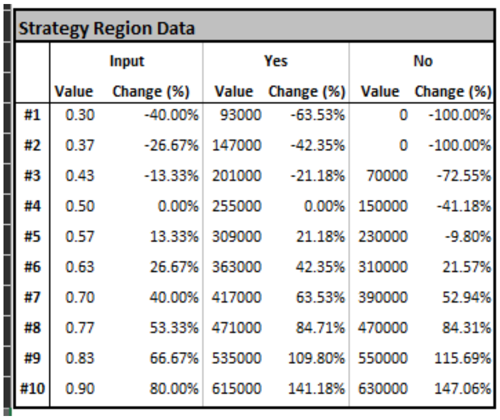
From the above question we can say that the cost of Information is $120,000 . So, the optimal decision tree doesn't change until the Survey Cost is increased to $120,000. The company can pay upto $120,000.

c) (5 pts) Ignore part b) and let C= $15,000. Conduct a sensitivity analysis on p: between 0.3 and 0.9 with 10 steps. Attached the strategy graph (Exhibit B) and paste the copy on your word document.



## PrecisionTree Sensitivity Analysis - Strategy Region

**Performed By:** Rahil
**Date:** Friday, December 10, 2021 5:11:21 PM
**Output:** Decision Tree 'PART A' (Expected Value of Entire Model)
**Input:** Prior probability of product success (B5)



Strategy Region Data

| | Input | | Yes | | No | |
|---|---|---|---|---|---|---|
| | Value | Change (%) | Value | Change (%) | Value | Change (%) |
| #1 | 0.30 | -40.00% | 93000 | -63.53% | 0 | -100.00% |
| #2 | 0.37 | -26.67% | 147000 | -42.35% | 0 | -100.00% |
| #3 | 0.43 | -13.33% | 201000 | -21.18% | 70000 | -72.55% |
| #4 | 0.50 | 0.00% | 255000 | 0.00% | 150000 | -41.18% |
| #5 | 0.57 | 13.33% | 309000 | 21.18% | 230000 | -9.80% |
| #6 | 0.63 | 26.67% | 363000 | 42.35% | 310000 | 21.57% |
| #7 | 0.70 | 40.00% | 417000 | 63.53% | 390000 | 52.94% |
| #8 | 0.77 | 53.33% | 471000 | 84.71% | 470000 | 84.31% |
| #9 | 0.83 | 66.67% | 535000 | 109.80% | 550000 | 115.69% |
| #10 | 0.90 | 80.00% | 615000 | 141.18% | 630000 | 147.06% |

- What is the approximate value of  p that changes the optimal strategy?

The optimal strategy changes at approximately 0.77. Beyond 0.77, the decision changes to not to  conduct the survey.

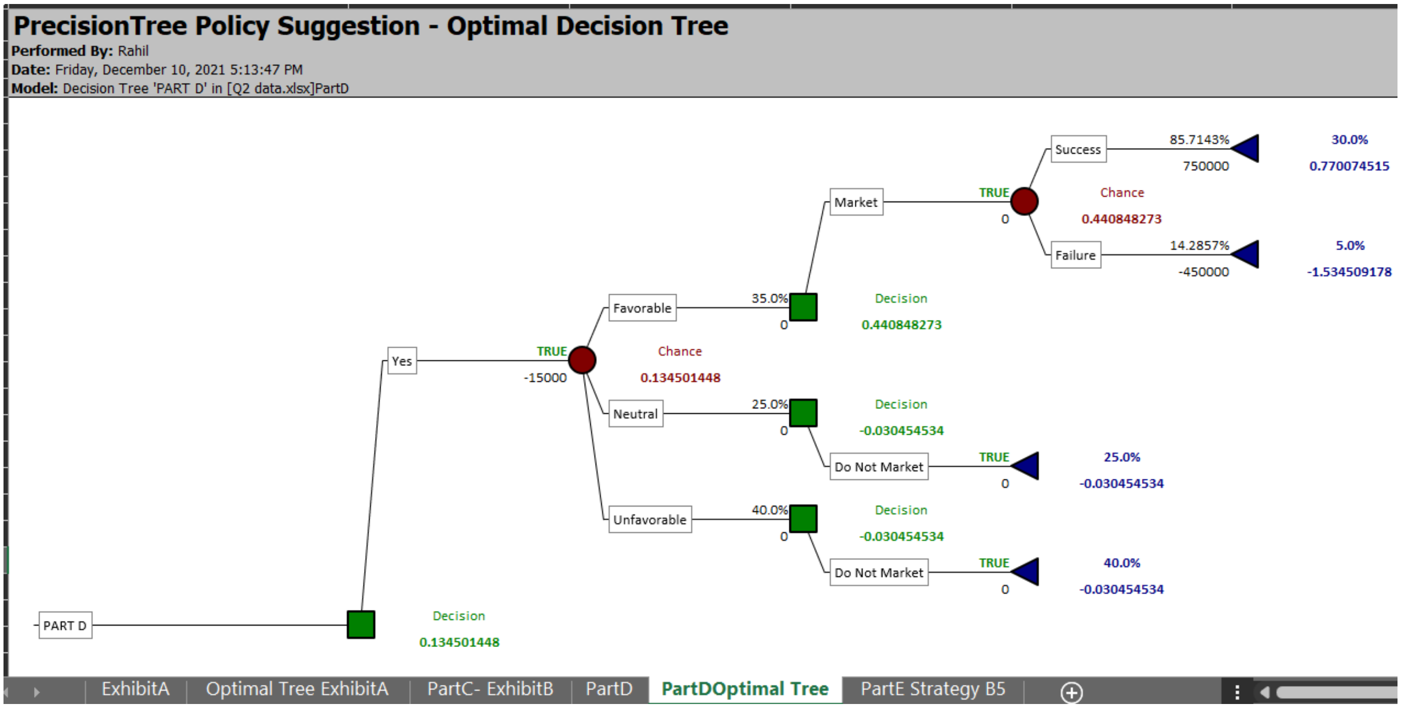
- Explain the results in detail.

In this scenario, p is the prior probability of success.

Higher probability of p will increase the expected value of the product without information. Once it reaches a high level the need to spend money for the survey will not offset its cost. The company's optimal strategy can make the second stage decision without the survey.

Now, we would like to find the strategy that maximizes the company’s expected utility with the risk tolerance R= 500,000.

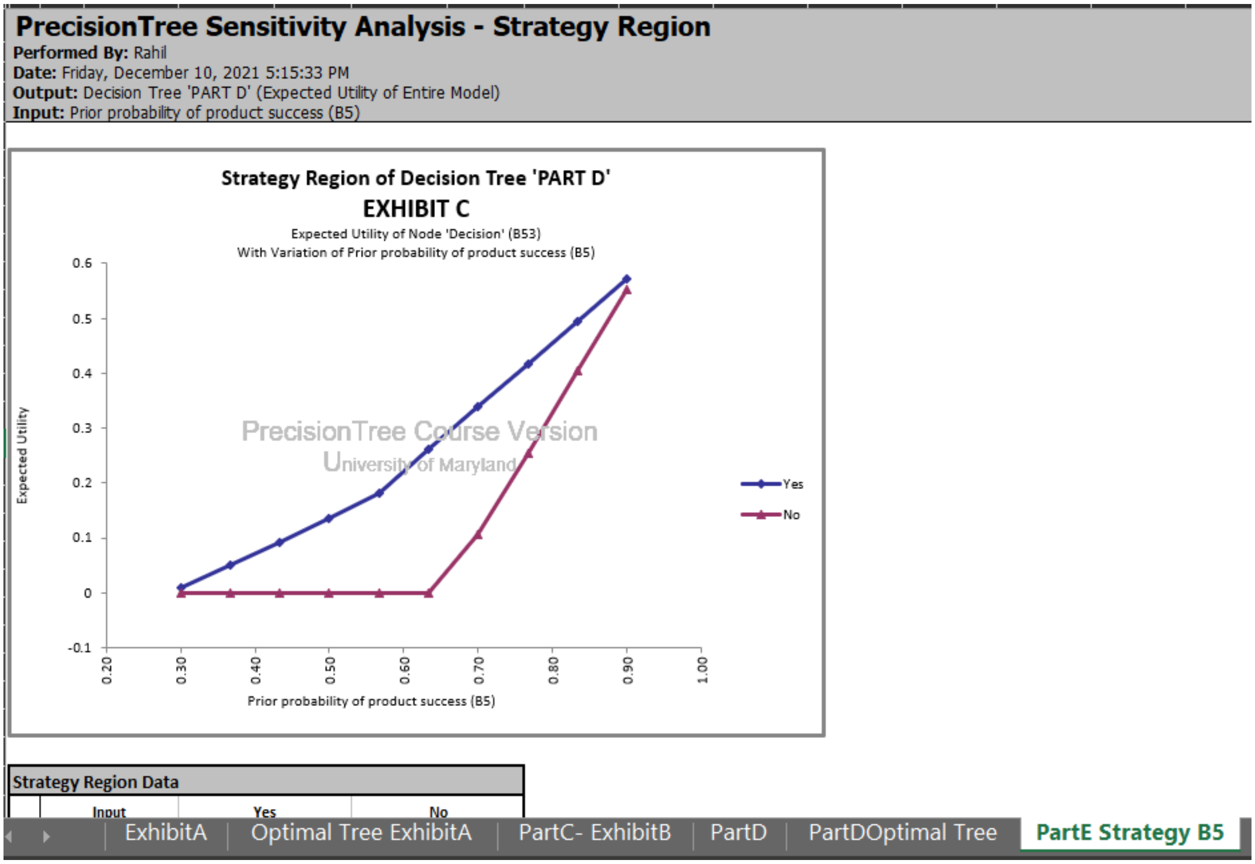
d) (5 pts) Generate the optimal decision strategy tree and paste the copy on your word document. Does this change the company’s decision?



**PrecisionTree Policy Suggestion - Optimal Decision Tree**
**Performed By:** Rahil
**Date:** Friday, December 10, 2021 5:13:47 PM
**Model:** Decision Tree 'PART D' in [Q2 data.xlsx]PartD

No, Adding Risk tolerance of R = 500,000 doesn’t change the company’s optimal decision strategy to take the survey.

e) (5 pts) Conduct a sensitivity analysis on p: between 0.3 and 0.9 with 10 steps. Attached the strategy graph (Exhibit C) and paste the copy on your word document. Explain the results. Particularly, explain how the second stage decision changes as p increases.



**PrecisionTree Sensitivity Analysis - Strategy Region**
**Performed By:** Rahil
**Date:** Friday, December 10, 2021 5:15:33 PM
**Output:** Decision Tree 'PART D' (Expected Utility of Entire Model)
**Input:** Prior probability of product success (B5)

Strategy Region of Decision Tree 'PART D'

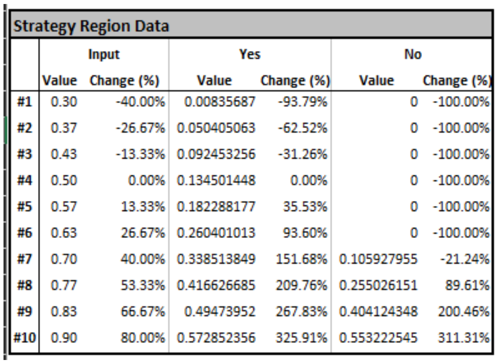
EXHIBIT C

Expected Utility of Node 'Decision' (B53)
With Variation of Prior probability of product success (B5)

Strategy Region Data

Input Yes No

ExhibitA
Optimal Tree ExhibitA
PartC- ExhibitB
PartD
PartDOptimal Tree
PartE Strategy B5



Strategy Region Data

| | Input | | Yes | | No | |
|---|---|---|---|---|---|---|
| | Value | Change (%) | Value | Change (%) | Value | Change (%) |
| #1 | 0.30 | -40.00% | 0.00835687 | -93.79% | 0 | -100.00% |
| #2 | 0.37 | -26.67% | 0.050405063 | -62.52% | 0 | -100.00% |
| #3 | 0.43 | -13.33% | 0.092453256 | -31.26% | 0 | -100.00% |
| #4 | 0.50 | 0.00% | 0.134501448 | 0.00% | 0 | -100.00% |
| #5 | 0.57 | 13.33% | 0.182288177 | 35.53% | 0 | -100.00% |
| #6 | 0.63 | 26.67% | 0.260401013 | 93.60% | 0 | -100.00% |
| #7 | 0.70 | 40.00% | 0.338513849 | 151.68% | 0.105927955 | -21.24% |
| #8 | 0.77 | 53.33% | 0.416626685 | 209.76% | 0.255026151 | 89.61% |
| #9 | 0.83 | 66.67% | 0.49473952 | 267.83% | 0.404124348 | 200.46% |
| #10 | 0.90 | 80.00% | 0.572852356 | 325.91% | 0.553222545 | 311.31% |

The second stage decision (marketing the product) is independent of the first stage decision (doing the survey). If the probability p, is too low, we will not market the product. At around 63% probability of the product success, the utility of taking the second stage decision without conducting the survey starts to increase at a higher rate than conducting the survey. When the probability is too  high, the likelihood of the product succeeding and its utility is enough to take the second stage decision to market the product without conducting a survey. The decision to shift to marketing of the product without a survey is higher than 90%.