Assignment 3: Promotions Management

Group 11: Christina Wang, Kailin Fu, Shun Guan, Sylvia Lu, Yiran Huang

April 20, 2020

Contents

1 Promotional event planning

2 Estimating lift factors and promotion ROI analysis

In this part of the assignment, we analyze the effectiveness and ROI of different promotions for Hellman's 32 oz Mayonnaise. The analysis is based on account level data at Jewel-Osco and Dominick's Finer Foods in Chicago. Use the table (data frame) hellmans_df in the file Hellmans.RData. hellmans_DF contains the following variables:

• account • product • week • units • dollars • feature_pctacv • display_pctacv

2.1 Question

Create a price variable for Hellman's 32oz mayo. Then, although not strictly necessary (because the estimated coefficients will scale in a linear regression), you should divide the feature and display columns (variables) by 100. Examine the feature and display variables. Provide summary statistics (number of observations, mean, standard deviation) and histograms of these variables, separately for both accounts. To what extent do these two promotional instruments differ? Calculate the correlations between feature_pctacv, display_pctacv, and price (use the cor function in R). Comment on your findings. Do the correlations indicate a potential problem for your regression analysis to be performed below?

```
hellmans_df$price = hellmans_df$dollars / hellmans_df$units
hellmans_df$feature = hellmans_df$feature_pctacv / 100
hellmans_df$display = hellmans_df$display_pctacv / 100

my_summary <- function(df, account) {
  df_local = df[df$account == account,]
    list("Feature Summary" =
       list("count" = length(df_local$feature),
            "mean" = mean(df_local$feature),
            "sd" = sd(df_local$feature)),
       "Display Summary" =
       list("count" = length(df_local$display),
            "mean" = mean(df_local$display),
            "sd" = sd(df_local$display)))
}

D_summary = my_summary(hellmans_df, "Dominicks")
J_summary = my_summary(hellmans_df, "Jewel")
print(D_summary)
```

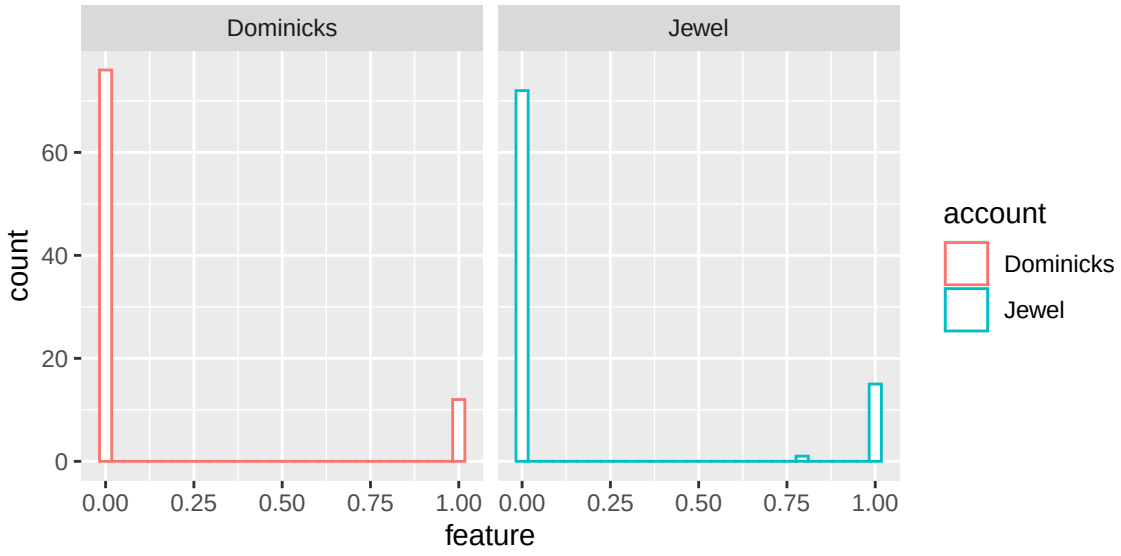
```
$`Feature Summary`
$`Feature Summary`$count
[1] 88

$`Feature Summary`$mean
[1] 0.1363636

$`Feature Summary`$sd
[1] 0.3451409


$`Display Summary`
$`Display Summary`$count
[1] 88

$`Display Summary`$mean
```

```
[1] 0.1206818

$`Display Summary`$sd
[1] 0.1762952
```

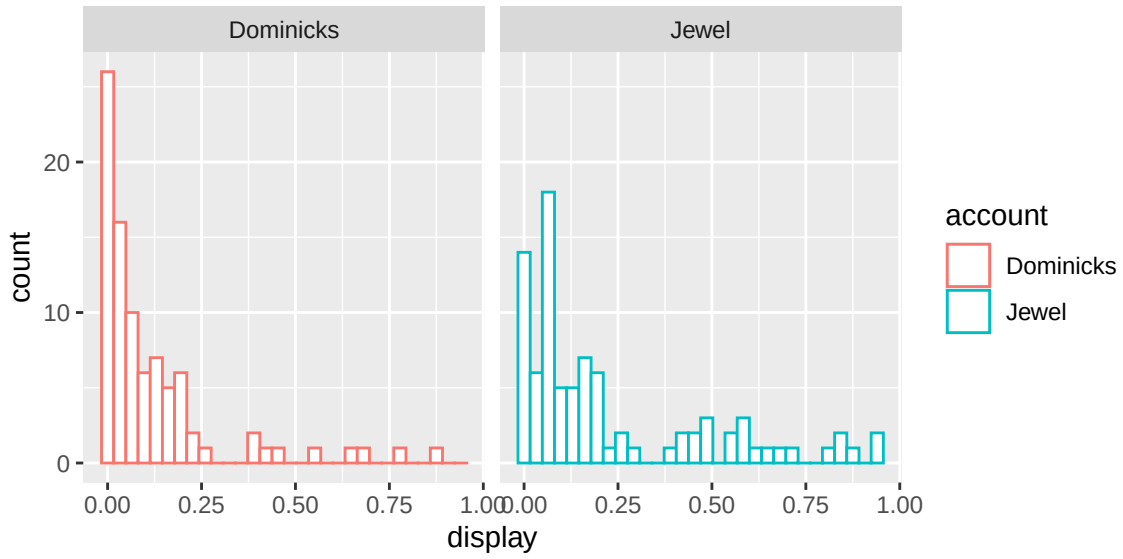
```
print(J_summary)
```

```
$`Feature Summary`
$`Feature Summary`$count
[1] 88

$`Feature Summary`$mean
[1] 0.1794318

$`Feature Summary`$sd
[1] 0.3834338


$`Display Summary`
$`Display Summary`$count
[1] 88

$`Display Summary`$mean
[1] 0.2298864

$`Display Summary`$sd
[1] 0.2581655
```

```
hellmans_df %>%
ggplot(data = ., aes(x=feature, color=account)) + geom_histogram(fill="white") + facet_grid(cols = vars
```




```
hellmans_df %>%
ggplot(data = ., aes(x=display, color=account)) + geom_histogram(fill="white") + facet_grid(cols = vars
```




```
#Correlations
cor(hellmans_df$feature_pctacv, hellmans_df$display_pctacv)
```

```
[1] 0.7599992
```

```
cor(hellmans_df$feature_pctacv, hellmans_df$price)
```

```
[1] -0.5747241
```

```
cor(hellmans_df$display_pctacv, hellmans_df$price)
```

```
[1] -0.6700056
```

2.2 Question

Estimate the log-linear demand model separately for each account, using price as the only explanatory variable. Then add the feature and display variables. Comment on the difference between the two regressions in terms of goodness of fit, and the price elasticity estimates. Is the change in price elasticity estimates as expected? What is the reason for this change? Are the coefficient estimates similar for both accounts?

```
D_lm =
hellmans_df %>%
filter(account == "Dominicks") %>%
glm(log(units) ~ log(price), data = .)

summary(D_lm)
```

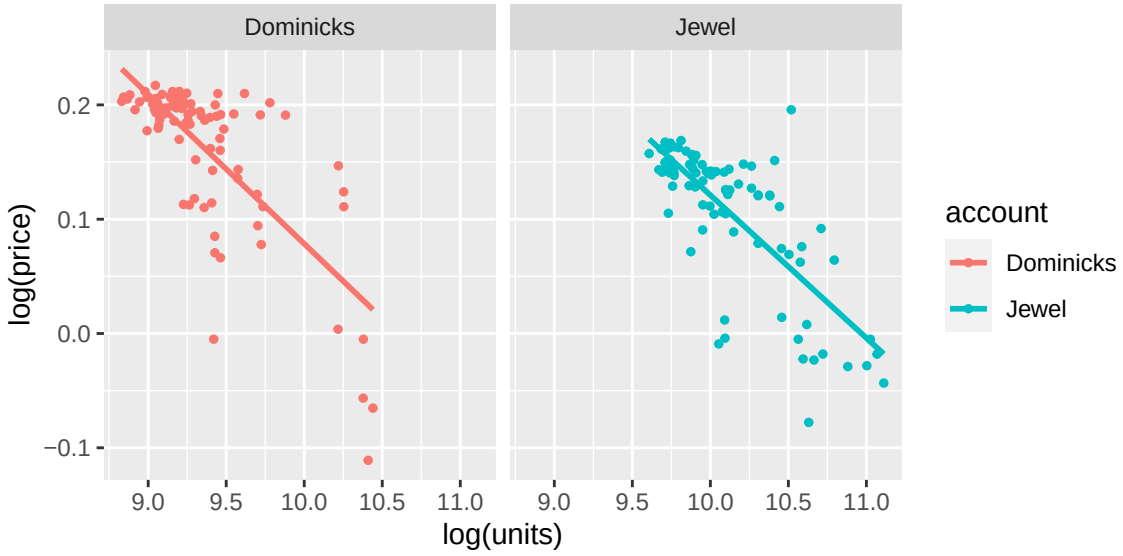
```

Call:
glm(formula = log(units) ~ log(price), data = .)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.63760  -0.17214  -0.01628   0.10558   0.79349

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  10.0368     0.0719  139.60  < 2e-16 ***
log(price)   -4.1665     0.4107  -10.15  2.3e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.06617719)

    Null deviance: 12.5037  on 87  degrees of freedom
Residual deviance:  5.6912  on 86  degrees of freedom
AIC: 14.753

Number of Fisher Scoring iterations: 2
```

```
J_lm =
hellmans_df %>%
filter(account == "Jewel") %>%
glm(log(units) ~ log(price), data = .)

summary(J_lm)
```

```

Call:
glm(formula = log(units) ~ log(price), data = .)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.59230  -0.16883  -0.03486   0.15152   0.81131

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.60443    0.05259  201.66   <2e-16 ***
log(price)  -4.58359    0.42660  -10.74   <2e-16 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.06161774)

    Null deviance: 12.4124  on 87  degrees of freedom
Residual deviance:  5.2991  on 86  degrees of freedom
AIC: 8.4712

Number of Fisher Scoring iterations: 2
```

```
#Two linear model demand-price graph
hellmans_df %>%
ggplot(data = ., aes(x= log(units), y = log(price), color=account)) + geom_point(size = 1, alpha = 1) +
  facet_grid(cols = vars(account)) + geom_smooth(method = "lm", se = FALSE)
```



2.3 Question

Consider the following three promotions:
(a) 15% TPR
(b) 15% TPR, 70% display
(c) 15% TPR, 70% display, 100% feature
Calculate the lift factors for each promotion for both accounts, based on the regression estimates in 2. Set estimates that are not statistically significant = 0.