BUAN 6392.0W1 Assignment 3

**Group 10**

Kushal Ayyagari Satyanarayana

Melissa Cunningham

Shubham Hota

Rohith Thotakura

Faraz Mohammed Mumtaz

1

1. Value of product video in three cases:

Prioritizing the introduction of product videos for popular (best-selling) products, new arrivals, products with price discounts, or products with other marketing promotions.

Value of product video = incremental sales from introducing video for popular products/new arrivals/discounted products/other marketing promotions.

Introducing product videos at an opportune time, such as during the holiday season.

Value of product video = incremental sales from introducing video per holiday season.

Introducing product videos at an opportune time based on some product-specific events, such as if a celebrity wore a dress on the Oscars night.

Value of product video = incremental sales from introducing video based on product-specific events.

b)

None of the expressions provide a correct estimate of the value of product video because they are all based on assumptions without any data to support them can lead to inaccurate results and there isn’t any randomized experiment performed to ensure the correct estimate or value of the product video.

2.

The experiment is designed to measure the causal effect of product videos on product sales on the retailer's website. The experiment involves creating short 15-20 second videos for 66 principal products (42 tops, eight bottoms, and 16 dresses), and randomly introducing these videos in three phases: 25 videos on February 17, 201X, 30 videos on March 23, 201X, and 11 videos on May 18, 201X. After hosting them on the website for over two months, the team removed these videos in three phases: on May 4, 201X, June 6, 201X, and July 18, 201X. The experiment was run during the 28 weeks of spring collection sales on the firm’s website from January 13, 201X to July 26, 201X.

The team collected weekly sales in numbers for all products during the experiment period. The team also collected information on other promotions run by the firm during the study period that could influence product sales, such as non-price promotions like catalog promotions, email promotions, and home page and category front page promotions, and price promotions like across-the-board price promotions and product-specific price promotions. The team ensured that there were no product reviews available on the retailer's website, and the possibility of other factors confounding the results of the experiment was absent.

We expect that: (1) the allocation of video in the full sample is proportionately transmitted to the subcategories of products (tops, dresses, bottoms, and auxiliary products) and (2) the mean characteristics of the treated and untreated products are statistically similar.

Based the data provided, we made a pivot table of the random check tab. Columns are Vid (treatment), rows are prod category and values are count of prodID. Then we calculated the percentage of # of treatment / total to get the proportion for each group. We further calculated a 2nd pivot table with rows as Vid, values as average ProdPrice to show that the average price between treatment/non treatment is about the same.

(2) Based on the t-test:

Text

Description automatically generated

Table

Description automatically generated

Using the excel tab ‘Random Check’, we created some pivot tables to analyze the data, shown below. For the product categories, the proportion of products that received video treatments are 13.3% for category 2 (bottoms), 18.2% for category 3 (dresses), and 20.6% for category 5 (tops). We calculated the 95% confidence interval for the percentage of treated products (based on p = 19.5% for all categories), and we found that the treatment proportion for each product category does fall within the confidence interval, meaning we have proportional assignment of video treatment across categories.   
  
  
  
Next, we need to check the summary statistics for the treatment and non-treatment products. Table shown below. We can see that the average product price, standard deviation, min and max product prices are all similar. This meets the criteria that the 2 groups are similar. Note that treatment group = 1 is video and treatment group = 0 is no video.   
  


3.)

1. Treatment Effect of Videos on Focal Product Sales

Difference in means estimate: The treatment effect is measuring the average weekly sales, and the treatment group is where the VidWk = 1, meaning a video is shown on a focal product during that week. The diff-in-means for the two groups is 22.49. We used a paired t-test to account for time seasonality changes between weeks. The results are statistically significant at alpha = 0.05 which means that we can reject the null hypothesis that there is no treatment effect. There is evidence that the effect of product videos will increase average weekly focal product sales by 22.49 with a 95% confidence interval of [8.59,36.39].

Paired t-test

data: vid\_group$meansales and novid\_group$meansales

t = 3.3203, df = 27, p-value = 0.002585

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

8.591839 36.388259

sample estimates:

mean of the differences

22.49005

Regression-based estimate: The regression equation has weekly focal product sales as the dependent variable and the video treatment + other promotions as independent variables. The regression equation and output are shown below in the text boxes. All the coefficients are statistically significant except for the price discount week variable. The interpretation of the effect of the product video is adding a focal product video leads to an increase in weekly sales of 26.29 after accounting for other promotional activities. The 95% confidence interval for the product video variable is [16.29, 36.29].

#Regression FP Analysis

reg1 <- lm(Sales~as.factor(VidWk)+as.factor(PriceDiscWk)+as.factor(EmailWk)+as.factor(CatalogWk)+as.factor(HomePgWk)+as.factor(CatPgWk),data=hw3)

summary(reg1)

Call:

lm(formula = Sales ~ as.factor(VidWk) + as.factor(PriceDiscWk) +

as.factor(EmailWk) + as.factor(CatalogWk) + as.factor(HomePgWk) +

as.factor(CatPgWk), data = hw3)

Residuals:

Min 1Q Median 3Q Max

-311.17 -70.18 -37.18 28.38 2160.58

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 100.176 1.609 62.269 < 2e-16 \*\*\*

as.factor(VidWk)1 26.293 5.102 5.153 2.63e-07 \*\*\*

as.factor(PriceDiscWk)1 -3.506 6.482 -0.541 0.58862

as.factor(EmailWk)1 98.791 30.610 3.227 0.00126 \*\*

as.factor(CatalogWk)1 215.992 23.293 9.273 < 2e-16 \*\*\*

as.factor(HomePgWk)1 94.192 12.795 7.362 2.03e-13 \*\*\*

as.factor(CatPgWk)1 101.453 13.369 7.589 3.65e-14 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 122.5 on 6821 degrees of freedom

(9 observations deleted due to missingness)

Multiple R-squared: 0.0403, Adjusted R-squared: 0.03946

F-statistic: 47.74 on 6 and 6821 DF, p-value: < 2.2e-16

These two estimates are similar but not the same. The reason is the interference effect of other promotional activities. The regression model accounts for the catalog, home page, category page, email promotions and price discount weeks to get an estimated measurement of effect on sales by adding a focal video. The means estimate used a paired t-test to account for variation in individual weeks which increased the standard deviation of our mean difference and resulted in a wider confidence interval. In either case, both results lead to the same overall conclusion – there is evidence that focal product videos do increase weekly sales.

4.)Treatment Effect of Videos on Coordinating Product Sales

The first regression analysis we performed was only CpSales (coordinating product sales) ~ Focal Prod Video treatment. This resulted in a significant estimate with value of 25.27 for VidWk, which mean a focal product video will increase weekly coordinating product sales by 25.27. Next, to analyses if Focal Prod Video treatment had any impact in pre or post week we performed regression of CpSales against VidWk, PreVidWk and PostVidWk. But it was inconclusive as we got significant result for only VidWk and result where insignificant for pre and post VidWk.

As we left out many other factors that could influence CpSales. Next, we performed a regression with all the various promotional variables as independent variables. The formula and results of this regression are below.

The focal product video treatment has an estimate value of 21.36 and it is statistically significant. We can interpret this value as adding a focal product video will increase weekly coordinating product sales by approximately 21.36. The 95% confidence interval for this estimate is [12.46, 30.27]. From the output we can see that there are many variables that are not statistically significant. The other independent variables that are significant are focal product catalog week, coordinating product email week, and coordinating product home page or category page weeks.

After the earlier regression analysis on focal product sales and this analysis, we can conclude that videos on focal products do lead to increased sales for both focal products and coordinating products.

Call:

lm(formula = CpSales ~ as.factor(VidWk), data = cp\_df)

Residuals:

Min 1Q Median 3Q Max

-91.07 -52.80 -34.07 9.20 2313.93

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 65.804 1.874 35.117 < 2e-16 \*\*\*

as.factor(VidWk)1 25.269 4.583 5.513 3.71e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 117.3 on 4706 degrees of freedom

Multiple R-squared: 0.006418, Adjusted R-squared: 0.006207

F-statistic: 30.4 on 1 and 4706 DF, p-value: 3.708e-08

#Regression - CP Analysis

reg2 <- lm(CpSales~as.factor(VidWk),data=cp\_df)

summary(reg2)

reg3 <- lm(CpSales~as.factor(VidWk)+as.factor(PreVidWk)+as.factor(PostVidWk),data=cp\_df)

summary(reg3)

reg4 <- lm(CpSales~as.factor(VidWk)+as.factor(FpPriceDiscWk)+as.factor(FpEmailWk)+as.factor(FpCatalogWk)+as.factor(FpHomePgWk)+as.factor(CpPriceDiscWk)+as.factor(CpEmailWk)+as.factor(CpCatalogWk)+as.factor(CpHomePgWk)+as.factor(CpCatPgWk),data=cp\_df)

summary(reg4)

Call:

lm(formula = CpSales ~ as.factor(VidWk) + as.factor(FpPriceDiscWk) +

as.factor(FpEmailWk) + as.factor(FpCatalogWk) + as.factor(FpHomePgWk) +

as.factor(CpPriceDiscWk) + as.factor(CpEmailWk) + as.factor(CpCatalogWk) +

as.factor(CpHomePgWk) + as.factor(CpCatPgWk), data = cp\_df)

Residuals:

Min 1Q Median 3Q Max

-217.72 -50.48 -31.97 10.52 2320.16

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 63.476 1.989 31.915 < 2e-16 \*\*\*

as.factor(VidWk)1 21.361 4.543 4.702 2.65e-06 \*\*\*

as.factor(FpPriceDiscWk)1 -9.385 6.484 -1.447 0.147837

as.factor(FpEmailWk)1 -19.683 28.494 -0.691 0.489739

as.factor(FpCatalogWk)1 117.984 14.613 8.074 8.56e-16 \*\*\*

as.factor(FpHomePgWk)1 4.524 9.728 0.465 0.641906

as.factor(CpPriceDiscWk)1 1.364 5.985 0.228 0.819732

as.factor(CpEmailWk)1 139.813 33.282 4.201 2.71e-05 \*\*\*

as.factor(CpCatalogWk)1 16.211 27.339 0.593 0.553242

as.factor(CpHomePgWk)1 90.622 25.480 3.557 0.000379 \*\*\*

as.factor(CpCatPgWk)1 92.431 22.418 4.123 3.80e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 115.5 on 4697 degrees of freedom

Multiple R-squared: 0.03993, Adjusted R-squared: 0.03789

F-statistic: 19.54 on 10 and 4697 DF, p-value: < 2.2e-16

Call:

Call:

lm(formula = CpSales ~ as.factor(VidWk) + as.factor(PreVidWk) +

as.factor(PostVidWk), data = df)

Residuals:

Min 1Q Median 3Q Max

-91.07 -52.62 -34.37 9.70 2313.93

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 65.623 2.080 31.550 < 2e-16 \*\*\*

as.factor(VidWk)1 25.450 4.670 5.450 5.3e-08 \*\*\*

as.factor(PreVidWk)1 10.360 6.202 1.670 0.0949 .

as.factor(PostVidWk)1 -10.253 6.710 -1.528 0.1266

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 117.3 on 4704 degrees of freedom

Multiple R-squared: 0.007624, Adjusted R-squared: 0.006991

F-statistic: 12.05 on 3 and 4704 DF, p-value: 7.472e-08Z

5.)

Marketing-Mix Decision

To test if product videos have an incremental effect on sales when combined with other kinds of marketing promotions, we need to consider interaction variables in the linear regression model. We ran the model below which shows both the regression formula and output.

All of the coefficients in the model are statistically significant at alpha = 0.05 except for the interaction between videos & home page weeks with a p-value of 0.055. This means that the combination of videos and marketing promotions does effect sales. If a product has both a video and it is a price discount week, we can expect sales to increase by 16.65 - 13.72 + 50.63 = 53.56 compared to no video and not a discount week.

An interesting observation is the interaction between product videos and catalog weeks with a coefficient of -224.19. To correctly interpret this, we need to consider 3 cases: if a product is featured in the catalog but does not have a product video, weekly sales increase by 241.63. If a product has a video and it is not a catalog week, weekly sales increase by 16.65. However, the combined strategy of product videos and the catalog promotion negatively impacts sales with a coefficient of -$224.19. One reason for this could be that the products highlighted in the catalog are different from the products with videos.

The incremental effect of a product having a video and being featured on the home page increases weekly sales by 83.99 and the incremental effect of a product having a video and being featured on the category page increases weekly sales by 194.30.

#Regression - Marketing Mix

reg4 <- lm(Sales~as.factor(VidWk)+as.factor(PriceDiscWk)+as.factor(EmailWk)+

as.factor(CatalogWk)+as.factor(HomePgWk)+as.factor(CatPgWk)+

as.factor(VidWk)\*as.factor(PriceDiscWk)+as.factor(VidWk)\*as.factor(CatalogWk)+

as.factor(VidWk)\*as.factor(HomePgWk)+as.factor(VidWk)\*as.factor(CatPgWk),data=hw3)

summary(reg4)

Call:

lm(formula = Sales ~ as.factor(VidWk) + as.factor(PriceDiscWk) +

as.factor(EmailWk) + as.factor(CatalogWk) + as.factor(HomePgWk) +

as.factor(CatPgWk) + as.factor(VidWk) \* as.factor(PriceDiscWk) +

as.factor(VidWk) \* as.factor(CatalogWk) + as.factor(VidWk) \*

as.factor(HomePgWk) + as.factor(VidWk) \* as.factor(CatPgWk),

data = hw3)

Residuals:

Min 1Q Median 3Q Max

-337.87 -70.25 -36.25 28.75 1988.42

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 101.247 1.611 62.866 < 2e-16 \*\*\*

as.factor(VidWk)1 16.650 5.369 3.101 0.001936 \*\*

as.factor(PriceDiscWk)1 -13.721 6.981 -1.966 0.049393 \*

as.factor(EmailWk)1 109.855 30.568 3.594 0.000328 \*\*\*

as.factor(CatalogWk)1 241.628 24.543 9.845 < 2e-16 \*\*\*

as.factor(HomePgWk)1 86.971 13.374 6.503 8.44e-11 \*\*\*

as.factor(CatPgWk)1 47.477 15.345 3.094 0.001983 \*\*

as.factor(VidWk)1:as.factor(PriceDiscWk)1 50.628 18.574 2.726 0.006432 \*\*

as.factor(VidWk)1:as.factor(CatalogWk)1 -224.192 74.705 -3.001 0.002701 \*\*

as.factor(VidWk)1:as.factor(HomePgWk)1 83.994 43.834 1.916 0.055385 .

as.factor(VidWk)1:as.factor(CatPgWk)1 194.301 31.435 6.181 6.74e-10 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 121.9 on 6817 degrees of freedom

(9 observations deleted due to missingness)

Multiple R-squared: 0.05018, Adjusted R-squared: 0.04879

F-statistic: 36.01 on 10 and 6817 DF, p-value: < 2.2e-16