Project 1

Causal inference in marketing experiments

Kimberly Venable – Dan Lemaire – Mike Fiore –

Submission Date: June 26th, 2017

MKTG6600 – Marketing Analytics

Executive Summary 2

Experiment Design 2

Hypothesis Testing 2

Hypothesis test for any difference between categories 2

Hypothesis test for a difference between one category and average of other categories 3

Hypothesis test for a difference between one category and the “2nd place” category 3

Results 3

Interaction Effects 4

# Executive Summary

The purpose of this analysis was to determine factors that cause online shoppers to spend more on the company website. The phrasing of the discount and the color of the letters used were both found to be statistically significant. The ad should be phrased as “% off” rather than “-%” and the letters should be red for maximum benefit.

# Experiment Design

Three hundred customers were assigned at random to be shown a discount in one of six different presentations. The amount of money spent on the website during that visit was tracked to assess which, if any, of the presentations encouraged the most spending. The customers were shown one of six presentations as such:

* “store wide price promotions of 10% off on what you purchase today”
* “store wide price promotions of 10% off on what you purchase today”
* “store wide price promotions of 10% off on what you purchase today”
* “store wide price promotions -10% on what you purchase today”
* “store wide price promotions -10% on what you purchase today”
* “store wide price promotions -10% on what you purchase today”

The design of the experiment was 2x3, with 50 customers in each category as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Design** | Color: Red | Color: Green | Color: Black |
| 10% Off | *50 Customers* | *50 Customers* | *50 Customers* |
| -10% | *50 Customers* | *50 Customers* | *50 Customers* |

# Hypothesis Testing

Since red is a color often associated with passion and impulsiveness, we believed that red letters would encourage more spending. Additionally, we believed that the “minus sign” was too small to be noticed and that “10% Off” would have a greater effect. As a result, we hypothesized that the “10% Off” announcement in red would have greater mean spending among the customers randomly assigned to this group. To test this, first we tested for any difference in means between groups. Next we tested for a difference between the hypothesized best group and the rest of the groups combined. Finally, we tested for a different between the best group and the “next best group”.

## Hypothesis test for any difference between categories

The point estimate for the mean $ spent in each group is shown in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Black | Green | Red |
| Minus | *$75.87* | *$53.81* | *$68.75* |
| Off | *$61.13* | *$49.79* | *$84.89* |

As we suspected, the point estimate for the red “10% off” message has the highest point estimate of the mean, but we needed to check whether the differences were statistically significant. An ANOVA test showed that the null hypothesis (mean spending in each group is the same) should be rejected at 95% confidence (p-value is effectively zero).

## Hypothesis test for a difference between one category and average of other categories

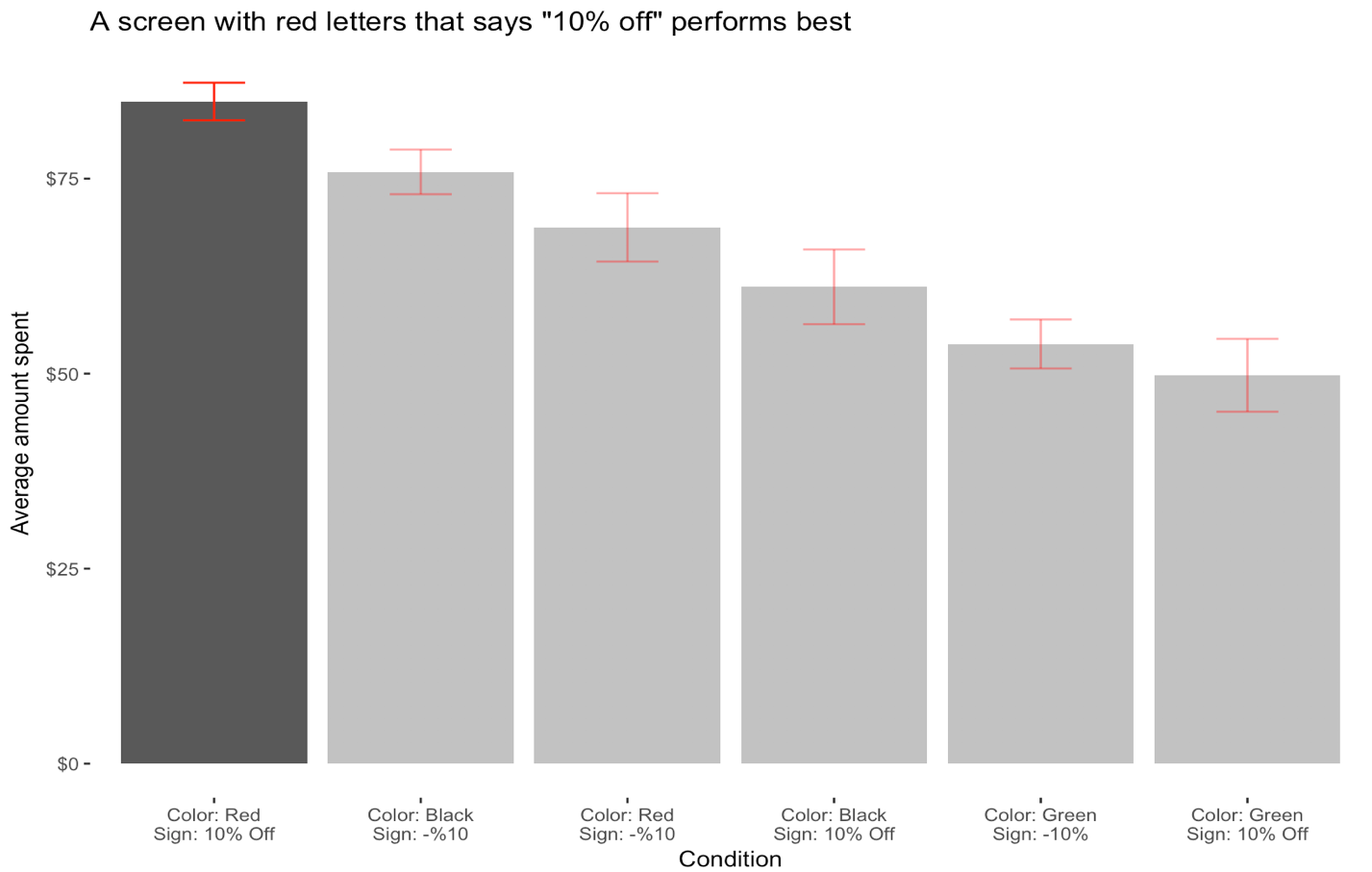
The data was split into 2 categories: one category contained all customers who saw a red “10% off” sign and the other category contained all other customers. The distributions of the two datasets were roughly normal but the variances were very different, so the Welch’s t-test was used. The null hypothesis (that the category of interest did not perform better than the others) was rejected. The category of interest performed better by over $20 on average.

## Hypothesis test for a difference between one category and the “2nd place” category

For the final hypothesis, we tested the best performing category against the second-best (a notice with black lettering that said “-10%”). For the same reasons above, we used a Welch’s t-test and found that the null hypothesis (that the category of interest did not perform better than the second-best category) was rejected. The red “10% off” notice performed over $9 better, on average, than the black “-10%” notice.

# Results

The company should prefer red lettering over black or green and should clarify discounts as “% off” rather than “-%” for the highest average spending. Our original hypothesis was clearly supported. The results can be seen in the visual below:



# Interaction Effects

Clear interaction effects are present. The impact of using “-10%” instead of “10% off” is different for each color. This means that the effect of each color and the effect of minus vs off do not explain how customers spend on their own… they require each other to fully explain the effect. A linear regression that does not account for interactions can explain only about 33% of the variance while a linear regression that includes interactions effects can explain about 45% of the variance. Tables that show the predicted group means with and without an interaction term are below.

**Predicted group means using Linear Regression without interaction term:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Black | Green | Red |
| Minus | *$68.93* | *$52.23* | *$77.26* |
| Off | *$68.06* | *$51.36* | *$76.38* |

**Predicted group means using Linear Regression with interaction term: (Same as actual group means)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Black | Green | Red |
| Minus | *$75.87* | *$53.81* | *$68.75* |
| Off | *$61.13* | *$49.79* | *$84.89* |