Star Digital A/B Testing Case

2023-10-01

Executive Summary

To maximize the conversion rate while effectively allocating marketing resources, Star Digital is recommended to run online advertising on Sites 1 through 5.

Procedure

After obtaining the data from a two-month experiment run on Sites 1 through 6, I initially applied exploratory data analytics techniques to better understand the data set. Then, I analyzed the impacts of advertisements and impressions by running logistic regressions.

Results

1. Effectiveness of Online Advertising

After running a logistic regression on the impact of advertisement on influencing users' purchase decisions, I discovered that the users who were exposed to Star Digital advertisement have about 8% higher odds of purchasing subscription packages than those who saw a different, unrelated advertisement.

2. Frequency Effects Analysis

A logistic regression on the total number of impressions and purchasing decisions proved that a unit increase in impression increases the odds of purchasing by about 3%. Therefore, I concluded that the more times a user is exposed to advertisements, the likelihood of purchasing increases.

3. Differences in Advertising Site Groups

The advertisement was shown on two groups of websites (Sites 1 through 5 and Site 6). After analyzing the potential impact of the two website groups by running a logistic regression, I discovered that the impressions made on Sites 1 through 5 increased the odds of purchase by about 3.3% while the impressions on Site 6 increased the odds by about 2%.

Recommendations

Since it is proven that online advertising is effective, Star Digital should continue its online advertising campaign. Additionally, advertising to users on Sites 1 through 5 will have higher odds of purchasing, even though it costs $5 more than Site 6, I recommend Star Digital to focus advertising on Sites 1 through 5.

Initial Analysis of Data

First, to understand the data set, I applied exploratory data analytics techniques. Out of 45 million users collected over the two-month experiment period, the sample data set is consisted of 25,303 users.

```
count(data)
```

```
## # A tibble: 1 × 1
##       n
##   <int>
## 1 25303
```

The dependent variable in the data set is whether a user made the purchase or not, which is a binomial data. There are two independent variables in the data set. The first independent variable is whether a user was shown a Star Digital advertisement or a different, unrelated charity advertisement.

```
data %>% group_by(test) %>% summarize(n())
```

```
## # A tibble: 2 × 2
##    test `n()`
##   <dbl> <int>
## 1     0  2656
## 2     1 22647
```

Test of 0 means that the user was not shown the Star Digital advertisement, while 1 means yes. Since only 10% of users were assigned to see non-Star Digital ad, it makes sense that there is much less number of users in that group compared to the users who saw the Star Digital ad.

Analyzing Experimental Design

Power and Sample Size

```
power.t.test(delta = 0.1, sd = 1, power = 0.9, sig.level = 0.05,
             type = "two.sample", alternative = 'two.sided')
```

```
##
##      Two-sample t test power calculation
##
##               n = 2102.445
##           delta = 0.1
##              sd = 1
##       sig.level = 0.05
##           power = 0.9
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

According to the power test, to observe the desired difference, there needs to be at least 2102 records in each group of control and treatment.

Since the collected data is over 45 million records, the experiment is likely to be overpowered, which means that the sample size is greater than the minimum required to examine desired impacts. An overpowered experiment is appropriate when a higher accuracy is required and a small difference between two groups is to be observed. As the conversion rate is small in nature, the overpowered experiment makes sense.

Checking the Validity of Random Assignment

The first step of the analysis involved confirming the validity of the random assignment of users into test and control groups. To do so, I compare the mean of total impressions for the test and control groups. If the mean is similar, then it is safe to assume that the users in both groups were exposed the similar amount of times to either the Star Digital ad or to the charity ad.

```
data = data %>% mutate(total_imp = imp_1 + imp_2 + imp_3 + imp_4 + imp_5 + imp_6)
data %>% group_by(test) %>% summarize(mean_total_impression = mean(total_imp))
```

```
## # A tibble: 2 × 2
##    test mean_total_impression
##   <dbl>                 <dbl>
## 1     0                  7.93
## 2     1                  7.87
```

After examining the mean of total impressions, the control group, which is represented as 0, and the test group, which is represented as 1, were similar. To confirm that these two means are statistically insignificant, I ran a t-test to compare the means of the two groups.

```
t.test(total_imp ~ test, data)
```

```
##
##  Welch Two Sample t-test
##
## data:  total_imp by test
## t = 0.12734, df = 3204.4, p-value = 0.8987
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.8658621  0.9861407
## sample estimates:
## mean in group 0 mean in group 1
##        7.929217        7.869078
```

The t-test yielded a p-value of 0.89, which is greater than the significance level of 0.05. Therefore, we can conclude that the two groups' mean impressions are similar.

Even though I used the mean total impressions to check the validity of the random user assignment, there is a concern as I was not able to examine other characteristics of the users. For example, a user's purchase decision can be influenced by his or her age, and I could not examine if the random assignment addressed these issues.

SUTVA Violation

In this experiment design, there is less likelihood of SUTVA violation since the users were only exposed to their assigned ads regardless of the six websites they visited as they were managed by a single ad serving network.

Star Digital Online Advertising Increases the Odds of Purchasing by about 8%

To examine whether being exposed to a Star Digital ad increase the probability of purchasing, I ran a logistic regression.

```
summary(glm(purchase ~ test, data, family='binomial'))
```

```
##
## Call:
## glm(formula = purchase ~ test, family = "binomial", data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.186  -1.186   1.169   1.169   1.202
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.05724    0.03882  -1.474   0.1404
## test         0.07676    0.04104   1.871   0.0614 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 35077  on 25302  degrees of freedom
## Residual deviance: 35073  on 25301  degrees of freedom
## AIC: 35077
##
## Number of Fisher Scoring iterations: 3
```

The logistic regression produced a coefficient of 0.07676 for the test variable. Since the dependent variable has been log transformed, I took the exponent to interpret the coefficient.

```
(exp(0.07676) - 1)*100
```

```
## [1] 7.97829
```

After taking the exponent, the output is about 7.98. This means that the odds of test group users purchasing is 7.89% higher than those in the control group. Therefore, I can conclude that the online advertising is effective for Star Digital.

However, the result needs to be interpreted with caution as the p-value of 0.061 from the logistic regression was above the significance level of 0.05.

Higher Frequency Leads to Higher Odds of Purchase

Next, I examined the potential impact of the number of impressions on purchase decisions through a logistic regression.

```
summary(glm(purchase ~ total_imp, data, family='binomial'))
```

```
##
## Call:
## glm(formula = purchase ~ total_imp, family = "binomial", data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.7410  -1.1265   0.1408   1.2166   1.2418
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.179392   0.014583  -12.30   <2e-16 ***
## total_imp    0.029201   0.001294   22.56   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 35077  on 25302  degrees of freedom
## Residual deviance: 34212  on 25301  degrees of freedom
## AIC: 34216
##
## Number of Fisher Scoring iterations: 5
```

To interpret the coefficient of 0.029 for the total impression in a meaningful way, I took the exponent of the coefficient.

```
(exp(0.029201) - 1)*100
```

```
## [1] 2.963153
```

The output of 2.96 can be interpreted there is a 2.96% higher odds of purchasing per a unit increase in impression. Therefore, it can be concluded that the more times a user is exposed to an ad, the more likely they will make a purchase.

Sites 1 through 5 Have Higher Odds of Purchasing

Lastly, I analyzed the impact of the two groups of websites that the ads were displayed on. I first created a new column that contains the sum of impressions from sites 1 through 5.

```
data = data %>% mutate(total_imp_1_5 = imp_1 + imp_2 + imp_3 + imp_4 + imp_5)
```

Then, I ran a logistic regression to analyze the impact of the impressions from sites 1 through 5 on the purchasing decisions.

```
summary(glm(purchase ~ total_imp_1_5, data, family='binomial'))
```

```
##
## Call:
## glm(formula = purchase ~ total_imp_1_5, family = "binomial",
##     data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.0075  -1.1297   0.1269   1.2119   1.2399
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.145722   0.013946  -10.45   <2e-16 ***
## total_imp_1_5  0.032438   0.001461   22.21   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 35077  on 25302  degrees of freedom
## Residual deviance: 34218  on 25301  degrees of freedom
## AIC: 34222
##
## Number of Fisher Scoring iterations: 5
```

I took the exponent of the coefficient 0.032438 to interpret the coefficient.

```
(exp(0.032438) - 1)*100
```

```
## [1] 3.296985
```

The output of 3.29 can be interpreted that there is about 3.3% higher odds of purchasing per unit increase of impression from site 1 through 5.

Next, I ran the logistic regression to analyze the impact of impressions from site 6.

```
summary(glm(purchase ~ imp_6, data, family='binomial'))
```

```
##
## Call:
## glm(formula = purchase ~ imp_6, family = "binomial", data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.2187  -1.1764   0.8864   1.1868   1.1868
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.022103   0.013433  -1.645   0.0999 .
## imp_6        0.019834   0.002927   6.776 1.24e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 35077  on 25302  degrees of freedom
## Residual deviance: 35014  on 25301  degrees of freedom
## AIC: 35018
##
## Number of Fisher Scoring iterations: 4
```

I took the exponent of coefficient to make a meaningful interpretation.

```
(exp(0.019834) - 1)*100
```

```
## [1] 2.0032
```

The output of 2 can be interpreted as that there is about 2% higher odds of purchasing per unit increase of impression from site 6.

In conclusion, though advertising on Sites 1 through 5 cost $5 more per 1,000 impressions than Site 6, users who see ads on Site 1 through 5 have higher odds of purchasing. Therefore, I would recommend Star Digital to advertise on Sites 1 through 5.