

Star Digital

Assessing the Effectiveness of Display Advertising

Anthony Meyers, Danny Moncada, Jack Quick, Kevin Grady, Michael Brucek

March 18, 2020

Executive Summary

This report was created in response to managerial inquiries into the efficacy of current online advertising. More specifically, it was designed to evaluate the impact of ad frequency on purchases and guide marketing budget allocation.

The test design chosen by Star Digital randomly divided 25,303 subjects into two groups. One group, referred to as the “treatment group”, was exposed to advertisements while the other, the “control group”, was not. The goal was to observe differences in purchase behavior between the treatment and control group to evaluate the incremental impact of advertising. For the purposes of this review, “effective” was simply defined as any statistically significant increase in purchasing as a result of ad exposure.

The testing conducted by Star Digital resulted in reliable evidence that advertising is effective at an overall level. 50.5% of those in the treatment group were converted to a purchase compared to 48.6% of the control group. The validity of the overall result and the experimental design were evaluated at a power of 80% and a significance level of 90%. Further validation at increased significance or power can be conducted upon request. Additional details regarding the testing and validation processes can be found in the Experimental Design Assessment section.

Overall results also show that increasing the frequency of advertising increases the likelihood of a purchase. However, the effect of this conclusion should be received with caution. The positive effect of ad frequency on purchase is present across the full spectrum of the treatment group, yet the impact is more pronounced for subjects who saw to an abnormally high number of ads. Exposure to online advertising is a function of time spent online, meaning those who are online often will naturally see more ads. The average number of ads viewed by those in the treatment group was 7.9, yet some individuals were exposed to as many as 521 ads. These high-frequency viewers are referred to as “anomalous” because they saw at least three times more ads than the top 75% of treatment subjects. This variance in exposure was controlled for in the assessment of results and can be viewed in more detail below

Finally, Star Digital will increase revenue by moving more advertising to the block of websites 1 through 5. The testing showed a valid and meaningful amount of expected profit from sites 1 through 5, even when the anomalous individuals were accounted for. Website 6 does not possess a significant relationship with revenue when anomalous behavior is removed. As a result, it would be beneficial to explore a shift in advertising spend from site 6 to sites 1 through 5.

Exploratory Data Analysis & Key Observations

Data Transformations

The raw data provides impression data at the website level granularity. Websites 1 through 5 were aggregated into a single variable to align with the business question. A count of total impressions across all sites was also created for use in future analysis. With those created, summary data of key variables can be seen in the following table.

##	purchase	test	imp_6	sum1to5
##	Min. :0.0000	Min. :0.000	Min. : 0.000	Min. : 0.000
##	1st Qu.:0.0000	1st Qu.:1.000	1st Qu.: 0.000	1st Qu.: 0.000
##	Median :1.0000	Median :1.000	Median : 1.000	Median : 1.000
##	Mean :0.5029	Mean :0.895	Mean : 1.784	Mean : 6.092
##	3rd Qu.:1.0000	3rd Qu.:1.000	3rd Qu.: 2.000	3rd Qu.: 4.000
##	Max. :1.0000	Max. :1.000	Max. :404.000	Max. :397.000
##	total_impressions			
##	Min. : 1.000			
##	1st Qu.: 1.000			
##	Median : 2.000			
##	Mean : 7.875			
##	3rd Qu.: 6.000			
##	Max. :521.000			

Purchase

The mean purchase value of .50 is evidence that half of the data is from an individual who purchased. This balance exists evenly in both the test and control groups.

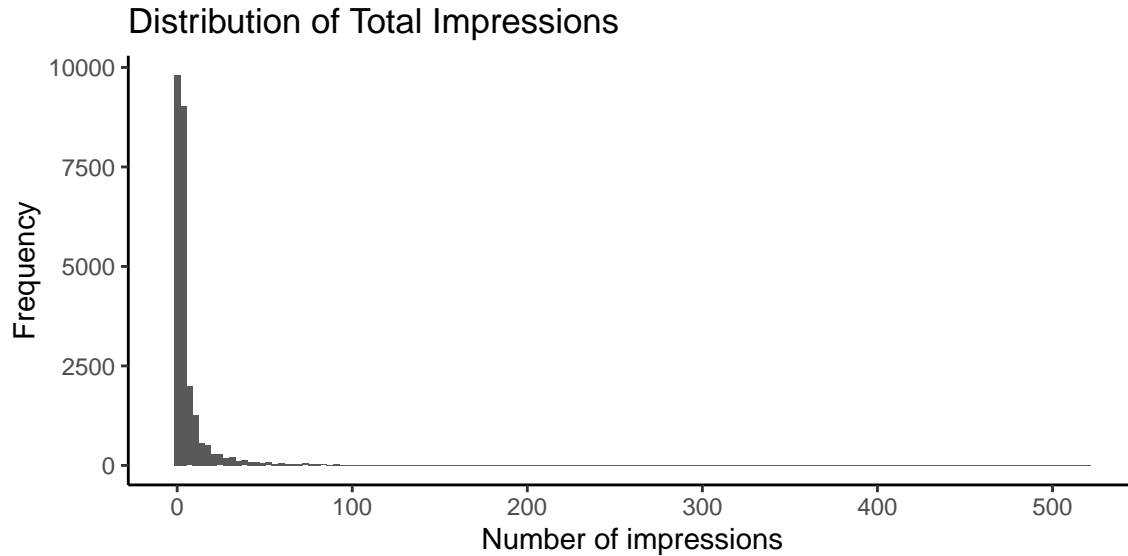
Test

The mean test value of .895 shows that nearly 90% of the data resides in the test, or “treatment”, group.

imp_6, sum1to5, and Total Impressions

The difference between the mean and max of these three variables is notable. This data skews heavily toward 0 with a lengthy tail of outliers stretching the range into hundreds of impressions.

The following table provides some perspective on how many advertisements consumers were exposed to. It's clear that the majority of people who saw ads were exposed to the mean number of about 8. However, the length of the tail highlights how severely the outliers skew the data. It will be important to control for excessive exposure in the analysis because their behavior and volume may impact key metrics. In addition, those who spend more time online are naturally more likely to purchase online.



Controlling for users who were exposed to an extreme number of advertisements requires the creation of an “anomalous” variable. A widely-accepted standard definition of anomalous is 1.5 times that of the third quarterly, but in an effort to provide a conservative analysis, three times the third quartile was how anomalous users were defined. A total of 1,942 individuals were classified as anomalous by this definition. 79.7% of these consumers were converted to a purchase, which is much more than their non-anomalous counterparts who only convert to purchases at a rate of 47.8%.

Finally, a quick review of correlation among the various websites shows that there is mild, if any, correlation between them. Validating this point clears the way to use sites 1 through 5 as a single variable without having to control for each site individually.

```
##           imp_1      imp_2      imp_3      imp_4      imp_5      imp_6
## imp_1 1.00000000 0.23301095 0.06121163 0.35228657 0.12940982 0.11331887
## imp_2 0.23301095 1.00000000 0.08288963 0.25635716 0.06384753 0.04969522
## imp_3 0.06121163 0.08288963 1.00000000 0.12602964 0.02886713 0.01992037
## imp_4 0.35228657 0.25635716 0.12602964 1.00000000 0.06943840 0.09649388
## imp_5 0.12940982 0.06384753 0.02886713 0.06943840 1.00000000 0.03184893
## imp_6 0.11331887 0.04969522 0.01992037 0.09649388 0.03184893 1.00000000
```

Experimental Design Assessment

This section addresses the viability of the experimental design to ensure results are reliable for inference.

Randomization

The first aspect under review is the randomization among the treatment and control groups. A T-Test was used to explore how the mean number of impressions and purchases compare between the test and control groups. A similar mean in each group implies effective randomization. Significantly different means suggests the attempt to randomize the data failed.

```
##
## Welch Two Sample t-test
```

```
##
## data: total_impressions by test
## t = 0.12734, df = 3204.4, p-value = 0.8987
## alternative hypothesis: true difference in means 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 mean number of impressions is very similar between the groups, and randomization for this variable is confirmed.

```
##
## Welch Two Sample t-test
##
## data: purchase by test
## t = -1.8713, df = 3309.2, p-value = 0.06139
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.039289257 0.000916332
## sample estimates:
## mean in group 0 mean in group 1
## 0.4856928 0.5048792
```

- The mean number of purchases is also very similar between the groups. Randomization for this variable is also confirmed.

Statistical Power

It is important to ensure that the experiment is powerful enough to detect statistically significant changes in purchase behavior between treatment and control groups. A power test was applied using the size of the control group as the value for “n” since 90% of the data was treated. The “power” parameter was set to a widely-accepted default value of .80, and the “significance level” parameter was set to .10 (90% significance). With the “delta” setting at “NULL”, this test will provide the level of change that the Star Digital experiment is capable of detecting.

```
##
## Two-sample t test power calculation
##
## n = 2656
## delta = 0.06824315
## sd = 1
## sig.level = 0.1
## power = 0.8
## alternative = two.sided
##
## NOTE: n is number in *each* group
```

The delta value of 0.06824315 implies that a change of up to 6.82% can be reliably detected in this experimental design. The difference between the treatment and control group in Star Digital’s is below this threshold, meaning this sample can be used to detect a causal effect of advertising on purchase.

Threats to Causal Inference

Threats to causal inference, such as omitted variables, should be noted. As always, unobserved confounds like device usage (mobile vs. tablet vs. desktop), customer age, and gender could impact purchase behavior but are not represented in the data. In addition, the sample has not been confirmed as representative of the population. There is potential for mismeasurement as offline purchases and offline ad exposure is not tracked. This raises mild measurement bias concern. Making a purchase is not likely to influence a user's decision to observe an advertisement, so there is little concern about simultaneity.

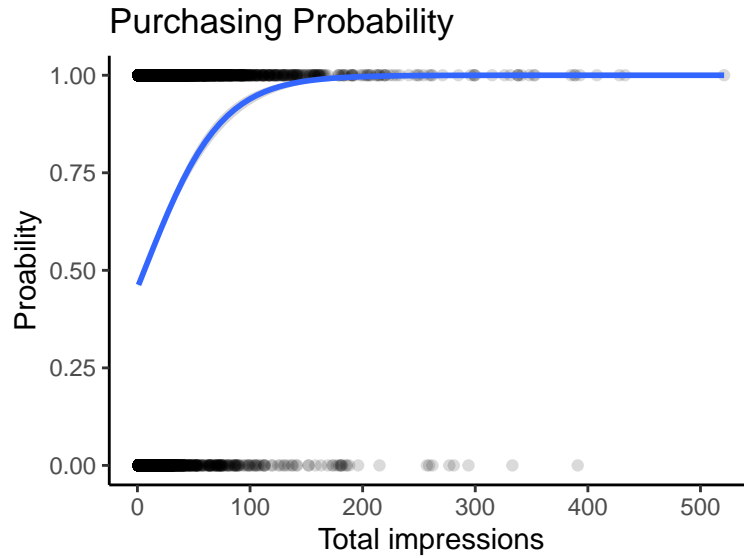
Results

The testing conducted by Star Digital resulted in reliable evidence that advertising is effective at an overall level. 50.5% of those in the treatment group were converted to a purchase compared to 48.6% of the control group. The experiment shows a valid, positive trend at the 90% significance when controlling for website and the anomalous class.

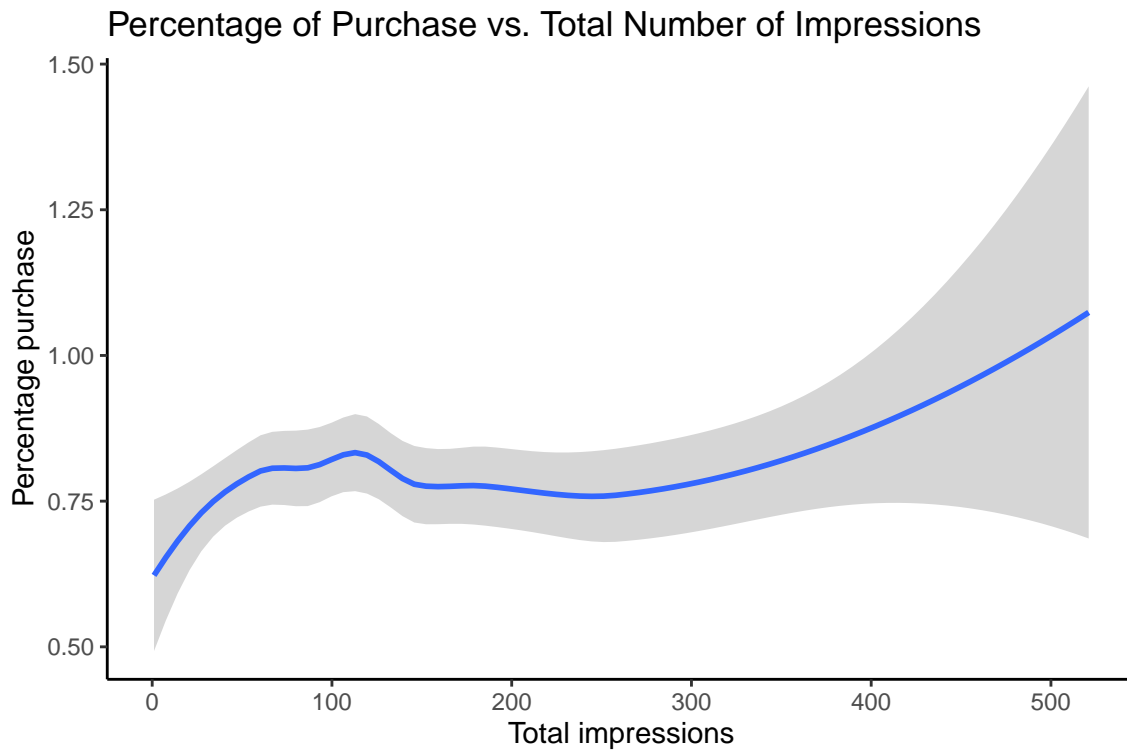
```
##
## Call:
## glm(formula = purchase ~ sum1to5 + imp_6 + test + anom, family = "binomial",
##      data = star_digital_df)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.2179  -1.1287   0.2126   1.2176   1.2643
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.208201   0.040177  -5.182 2.19e-07 ***
## sum1to5      0.021589   0.002108  10.243 < 2e-16 ***
## imp_6        0.006332   0.002828   2.239  0.0251 *
## test         0.071027   0.041733   1.702  0.0888 .
## anom         0.590890   0.094884   6.227 4.74e-10 ***
## ---
## 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: 34145  on 25298  degrees of freedom
## AIC: 34155
##
## Number of Fisher Scoring iterations: 5
```

It is important to remember that logistic regression model coefficients need to be interpreted with caution as they represent log-odd ratios and not direct probabilities. The positive coefficient value for the “test” variable implies the relationship between ad exposure and purchasing is positive (purchasing increases with exposure to advertisements). The .088 p-value is within the .100 significance threshold that was tested during the power test, meaning this result is accepted as a significant variable in the analysis.

In addition, the experiment provided insight into the effect that frequency of ad viewing has on a potential customer's likelihood to purchase. The following plot shows the distribution of customers who purchased as a function of the number of ads they were served. There is a clear and positive relationship between ads viewed and purchase probability: The more ads someone views, the more likely they are to purchase.

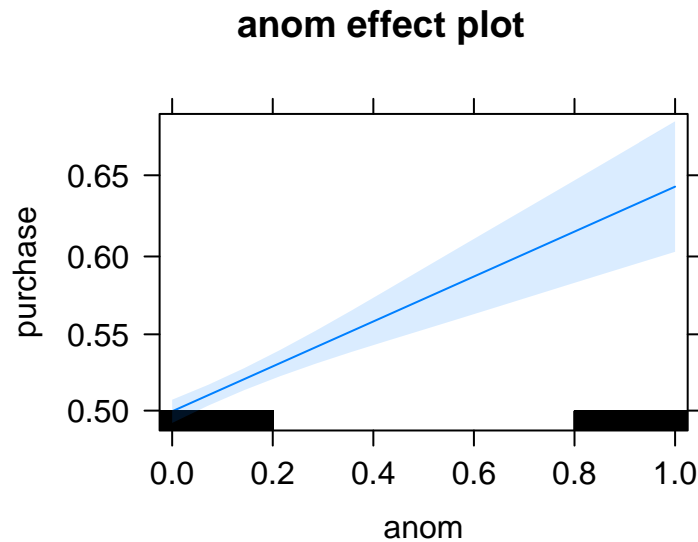


A secondary view of this same relationship is shown on the following plot. This visualization displays the percentage of customers who made a purchase for each number of ads they were served. The chart extrapolates the trend beyond the 100% purchase point as a result of the linear relationship between ad views and purchases.

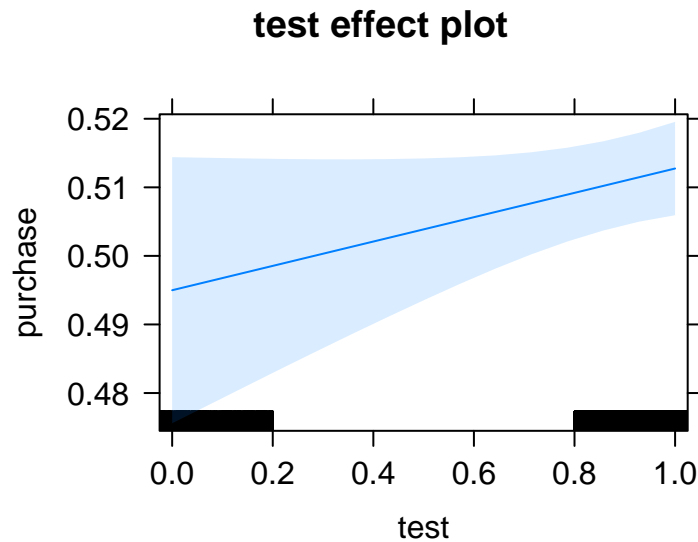


Again, a clear and positive relationship between ad frequency and purchasing is seen. This also shows the importance of controlling for anomalous exposure. The cutoff for anomalous views was 21 ads, after which the trend in purchasing remains high and relatively constant.

It was noted that anomalous customers purchase at nearly an 80% rate. The plot below shows the frequency effect of ads on anomalous viewers, and it is very pronounced.



It is important to control for this behavior in order to ensure the conclusion regarding frequency holds true across those who were not anomalous. The plot below shows the effect on the treatment group when the anomalous class has been controlled for.



Ad viewing frequency continues to effect purchasing in a positive manner even with the outliers removed, although the impact of that effect is much less.

Exploring which, if any, of the website groups was the most effective in terms of generating revenue required an expected value analysis. Impressions data was transformed into single unit costs (\$25 per 1,000 impressions for websites 1 through 5 and \$20 per 1,000 impressions for website 6). The purchase variable was transformed into its lifetime revenue value of \$1,200.

This transformation allows linear models to estimate an expected revenue value per 1 unit of advertising, relative to the cost of advertising one unit on each website. The output of this analysis is below. Note that this model continues to control for the websites and the anomalous user class.

```
##
## Call:
## lm(formula = revenue ~ as.factor(test) + as.factor(anom) + site1to5_cost +
##     site6_cost, data = star_digital_spend_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1653.97  -573.05    34.42   625.46   649.07
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      550.56      11.55  47.671 < 2e-16 ***
## as.factor(test)1    20.43       12.11   1.687  0.0916 .
## as.factor(anom)1   278.78      19.30  14.444 < 2e-16 ***
## site1to5_cost      82.27       10.13   8.119 4.91e-16 ***
## site6_cost        18.63       27.57   0.676  0.4992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 590.5 on 25298 degrees of freedom
## Multiple R-squared:  0.03142,    Adjusted R-squared:  0.03126
## F-statistic: 205.1 on 4 and 25298 DF,  p-value: < 2.2e-16
```

The output of the linear model shows that the coefficients for both websites 1 to 5 and website 6 are both positive, implying that users who saw impressions on these sites eventually do make purchases. But, the coefficients of the website variables should be compared (expected value revenue from 1 unit of advertising) to the cost of advertising one unit on these websites (\$25 per 1,000 impressions for websites 1 through 5 and \$20 per 1,000 impressions for website 6).

One unit of advertising, defined as 1,000 impressions, generates an expected revenue of \$82.27 for websites 1 to 5 and \$18.63 for website 6. Compared to the cost of the advertising, one unit of advertising does not generate any expected profit on website 6 (\$18.63 - \$20 = -\$1.37) where each unit of advertising on websites 1 to 5 generate an expected profit of \$57.27 (\$82.27 - \$25).

Star Digital should focus its advertising on websites 1 through 5 over website 6. However, the coefficient of website 6 in the linear model above is not statistically significant. This implies that caution should be used in making assumptions on advertising on website 6 until more data is available.

Conclusion / Summary

The three main takeaways of this report are as follows:

- The data collected and provided by Star Digital shows a significant positive effect as a result of digital advertising, that is to say being a part of the treatment group. The probability of a user making a purchase increases by approximately 51.77% when controlling for the quantity of advertising impressions the user receives and whether or not a user receives an anomalous number of impressions.
- Exposure to digital advertising is not the only factor that has a causal effect on the probability of a user making an online purchase. The total number (magnitude) of ad impressions that a user is exposed

to, whether for a Star Digital advertisement or for a “dummy” ad as the control group received, also has a direct impact on the probability of a purchase. As the total number of ad impressions served to a user increases, so does the probability of a purchase being made by that user. It may be beneficial to note that the frequency effect of advertising exposure is more apparent in anomalous users who were exposed to an extreme amount of ads. Exposure to online advertising is a function of time spent online, meaning those who are online often will naturally see more ads. Nevertheless, there is enough evidence present in the data to conclude that advertisement exposure has an effect on purchasing probability.

- Given the higher profitability of advertising on websites 1 through 5, Star Digital should consider reallocating current marketing spend away from site 6 in order to increase the allocation going to sites 1 through 5. When a user does not exhibit anomalous web browsing behavior, impressions on website 6 do not appear to have a significant relationship with revenue.