Digital ad campaigns effectiveness on customer purchase decisions

Objectives

Digital ad campaigns on the company's website with tracking cookies to record customer purchases are commonly used for measuring the effectiveness of the ad exposure treatment. This analysis aims to answer several questions relating to the marketing campaign's effectiveness by demonstrating the explanatory indicators, whether the marketing strategies should be kept or reconsidered, and inform the marketing manager about opportunities and obstacles in future experiments.

Methodology

Using the provided data set for this analysis, it seems that to measure the effectiveness of the advertisement campaign exposure on the treatment and control group as assigned, it is rational to measure the purchase of individual users by applying the treatment effect. There is another variable in the data set which is 'impressions' - the ad exposure frequency on each user could be a potential predictive variable.

Since the purchase variable is binary, with 0 as no purchase had been made, and 1 as purchased, the logistic regression can be used for predicting the outcomes. Focusing on the treatment effect application, we would like to assess the campaign effectiveness if it is a causal effect of the exposure treatment.

This data set contains 20,000 observations, representing 20000 users with different outcomes and random assignment on test and control groups in the ratio of 90% treatment / 10% control. The test group are chosen to see the ads, and the control does not see any ads at all.

For better formulation purposes, let's define the following variables on observations from 20,000 users::

- W: Treatment assignment (1 if in test group, 0 if in control group). $W_i \in \{0, 1\}$
- Y: Observed outcome (customer purchase decision). $Y_i \in \mathbb{R}$

There are two potential outcomes for each user:

- $Y_i(1)$: The response we would have measured if the i-th user received treatment ($W_i = 1$)
- $Y_i(0)$: The response we would have measured if the i-th user received no treatment ($W_i = 0$)

Identifying the two potential outcomes emphasizes the effect of the treatment on each individual is different between the outcomes that would have been observed with and without the treatment. They are 'potential' for the reason that they are usually unobserved.

The causal effect of the treatment can be defined as the difference between the two potential outcomes:

$$\tau_i = Y_i(1) - Y_i(0)$$

Data Visualization

```r
# Load necessary libraries:
library(sjPlot)
```

```
Learn more about sjPlot with 'browseVignettes("sjPlot")'.
```

```r
library(ggplot2)
```

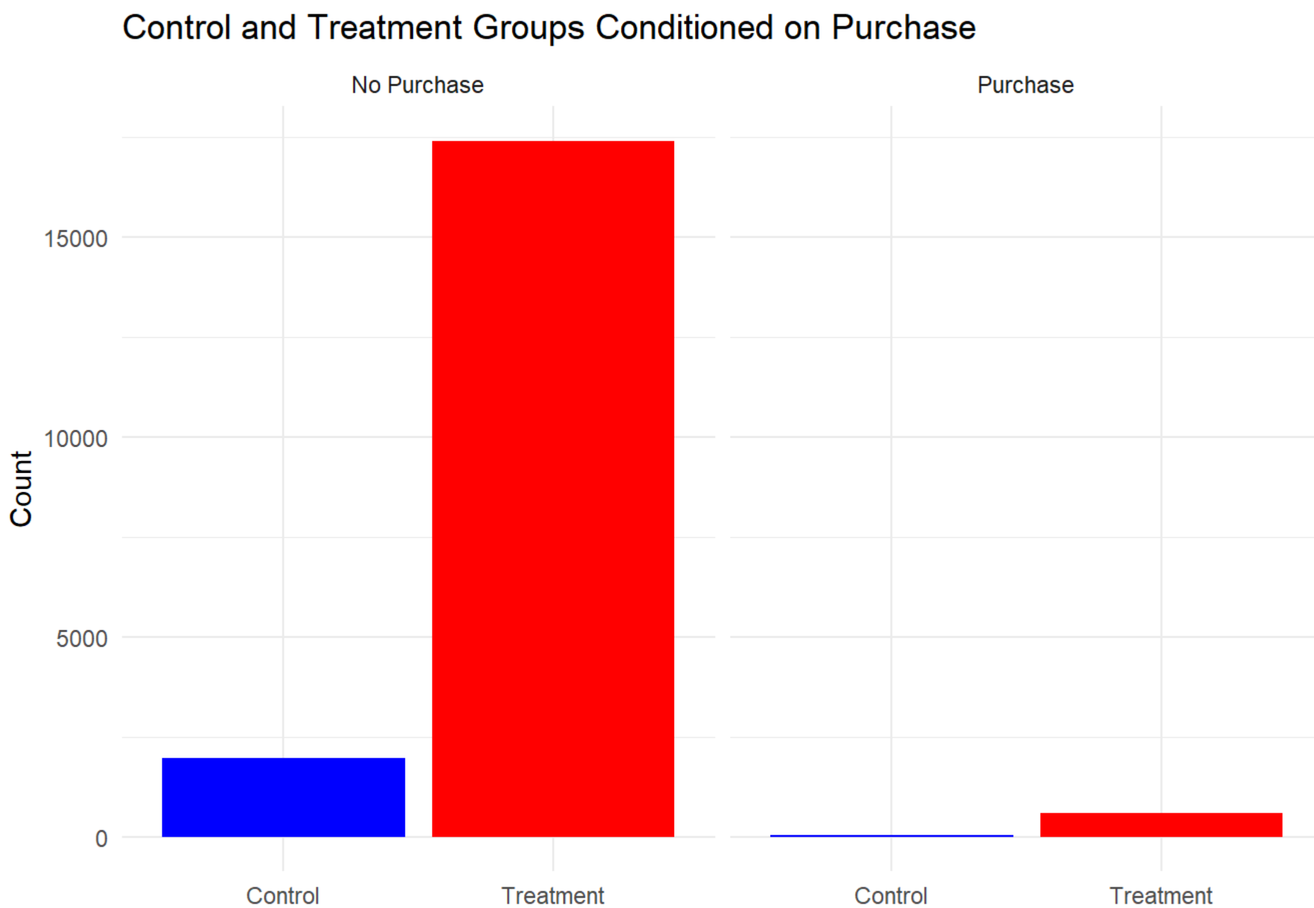
```
Warning: package 'ggplot2' was built under R version 4.3.3
```

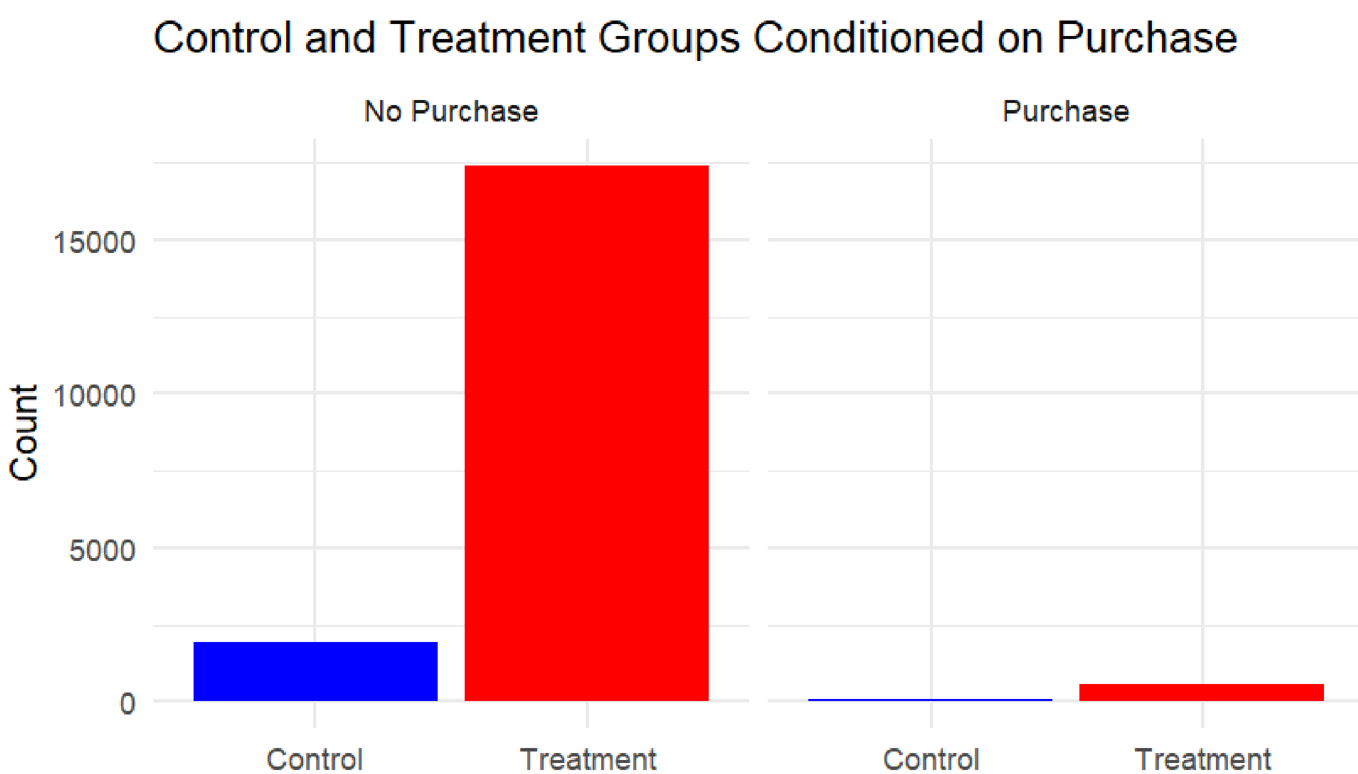
```r
# Load the data
data <- read.csv("CaseData.csv", sep = ";")
```

Make a bar plot to see the distribution of the control group and the treatment group in terms of making purchase decisions:

```r
# Convert the 'test' variable to a factor for clarity in the plot
data$group <- factor(data$test, levels = c(0, 1), labels = c("Control", "Treatment"))

# Create the plot
ggplot(data, aes(x = group)) +
  geom_bar(aes(fill = group), position = "dodge") +
  facet_wrap(~purchase, labeller = labeller(purchase = c('0' = "No Purchase", '1' = "Purchase")))
  scale_fill_manual(values = c("Control" = "blue", "Treatment" = "red")) +
  labs(title = "Control and Treatment Groups Conditioned on Purchase",
       x = "",
       y = "Count") +
  theme_minimal() +
  theme(legend.position = "none")
```





The above plot has shown the great imbalance between the control and treatment groups, purchase and non-purchase groups. In comparison, approximately 17,000 users in the treatment group had not made a purchase, and only about 600 users who had received the treatment made the payment. While there was no treatment in the control group, there were few purchases made by the users.

True Estimation

Implying the formulas for calculating the ATE (average treatment effect), ATET (average treatment effect on the treated), the baseline conversion rate and lift in conversion, on observed data:

```
# The true average treatment effect (ATT)

# Calculate the average outcome for those treated
mean_treated <- mean(data$purchase[data$test == 1])

# Calculate estimate of the counterfactual outcome for the same individuals had they not been trea

mean_counterfactual <- mean(data$purchase[data$test == 0])

# Calculate ATET

atet <- mean_treated - mean_counterfactual

# Baseline conversion rate on this group's purchase is the mean of counterfactual outcomes

bcr <- mean_counterfactual

# Conversion rate of the test group:
tcr <- mean_treated

# Print the results:
cat("Observed Baseline Conversion Rate:", round(bcr * 100, 2), "%\n")
```

```
Observed Baseline Conversion Rate: 2.49 %
```

```
cat("Observed Treatment Conversion Rate:", round(tcr * 100,2), "%\n")
```

```
Observed Treatment Conversion Rate: 3.23 %
```

```
# Calculate lift in conversion:
lift <- (tcr - bcr) * 100

cat("Observed Lift in conversion:", round(lift,2), "%\n")
```

```
Observed Lift in conversion: 0.74 %
```

The baseline conversion rate in this case is the total conversion rate of the control group, as the baseline conversion rate is the likelihood of purchasing without any intervention (treatment received), and the control group, in this case, was not exposed to the ad (did not receive any treatment). Therefore this indicator can be calculated directly from the observed data.

In the next step we imply the model considering the exposure assignment D, and investigate if there is one-sided non compliance situation:

```
#Create a binary 'impressions' variable from the data:

data$any_impressions <- ifelse(data$impressions > 0, 1, 0)

# Check if there is 0 values:
table(data$any_impressions)
```

```
    1 
20000 
```

The result of the binary impressions variable suggests no 0 value, with the values defined for the control group being the number of impressions that would have been shown to them if they had been treated, even tho there was no ad exposing to them. But let's look at the cross table of test group assignment $(test)$ and actual treatment received $(impressions)$:

```
W <- data$test
n <- 20000
table(W,data$any_impressions)/n
```

```
W        1
  0 0.1003
  1 0.8997
```

The cross table suggests that $D[W=1]=0.8997$ and $D[W=0]=0.1003$ for D always positive, the result suggests 89.97% of the data set belongs to the test group and has exposed to the ad, while 10.03% of entire data set belongs to the control group and would have exposed to the ad. Approximately 90% of the data set in the test group, actually received the treatment, indicating that the treatment group always receives the treatment, and 10% of the data set is the control group, who were not exposed to the ad and did not receive the treatment. In order words, the ratio (90% and 10%) that the company decided to divide the customer set into test and control groups, was implied precisely in terms of compliance absolute of receiving the treatment.

Regression

We run a logistic regression that estimates the effect of treatment on purchasing decisions by this equation $Y_i = \alpha + \beta W_i + \epsilon_i$. $purchase$ is assigned as the response variable, while $test$ serves as the predictor variable.

```r
# Run a linear regression on the data to estimate the treatment effect
model <- glm(purchase ~ test, data = data, family = binomial)

# Print the summary of the regression output
tab_model(model, show.stat = TRUE)
```

```
Profiled confidence intervals may take longer time to compute.
  Use `ci_method="wald"` for faster computation of CIs.
```

	purchase			
Predictors	*Odds Ratios*	*CI*	*Statistic*	*p*
(Intercept)	0.03	0.02 – 0.03	-25.60	**<0.001**
test	1.31	0.99 – 1.77	1.80	0.072
Observations	20000			
R^2 Tjur	0.000			

	purchase			
Predictors	*Odds Ratios*	*CI*	*Statistic*	*p*
(Intercept)	0.03	0.02 – 0.03	-25.60	**<0.001**
test	1.31	0.99 – 1.77	1.80	0.072
Observations	20000			
R^2 Tjur	0.000			

The odds ratio for the intercept is 0.03, meaning that the odds of making a purchase where there is no treatment effect ($test$ == 0) is very low, typically showing the base conversion rate is low (3%). The odds ratio for the $test$ is 1.31, meaning that the existence of treatment could increase the odds of purchasing to 31%.

However, this $test$ variable is not statistically significant since its p-value is higher than 0.05. Moreover, the R-squared has a very low value indicating that the purchase variable is not explained well by the treatment assignment effect.

Next, we figure out the baseline conversion rate and lift in conversion using the prediction from the logistic regression model.

The baseline conversion rate is the conversion rate observed from the control group, who has not been exposed to the ad. Represents the normal rate without any influence from the marketing campaign.

The lift in conversion rate indicates how much of the conversion rate has increased after the campaign, compared to the baseline conversion rate. Showing a conversion lift answers the question: “How much in conversion rate was increased caused directly by the ads campaign?”.

```
# Predicted probabilities for the treated group based on the logistic regression model:
prob_treated <- predict(model, type = 'response')

# Create a new dataset for the counterfactual scenario
data_cf <- data
# Set treatment to 0 for all
data_cf$test <- 0

# Predict probabilities for the counterfactual scenario (treated individuals as if not treated)
prob_cf <- predict(model, newdata = data_cf, type = "response")

# Calculate the treatment effect on the treated
atet_p <- mean(prob_treated[data$test == 1]) - mean(prob_treated[data$test == 0])

# The baseline conversion rate and lift in conversion from the prediction:

bcr_p <- mean(prob_cf[data$test == 1]) # considered in the counterfactual scenario

# The lift in conversion is the treatment effect on the treated
lift_in_conv <- atet_p

# Print the results:

cat("Observed Baseline Conversion Rate:", round(bcr * 100,2), "%\n")
```

```
Observed Baseline Conversion Rate: 2.49 %
```

```
cat("The estimated Baseline Conversion rate is", round(bcr_p *100, 2), "%\n")
```

```
The estimated Baseline Conversion rate is 2.49 %
```

```
cat("Observed Lift in conversion:", round(lift, 2), "%\n")
```

```
Observed Lift in conversion: 0.74 %
```

```
cat("The estimated Lift in Conversion:", round(lift_in_conv*100, 2), "%\n")
```

```
The estimated Lift in Conversion: 0.74 %
```

We have calculated the baseline conversion rate and lift in conversion, with the alignment in the estimated results and observed values. We can conclude:

- Baseline Conversion Rate: 2.49%
- Treatment Conversion Rate: 3.23%
- Lift in Conversion: 0.74%

Incremental lift (Incremental effect) - a measurement of the marketing campaign effectiveness

To evaluate the effectiveness of the marketing campaign, the incremental effect or incremental conversion rate is pretty useful for understanding the additional value, which is generated by the treatment effect.

Incremental lift conversion rate – this indicator gives an understanding of the additional/possible impact of the ad campaign beyond what would have occurred naturally. It is more into explaining the gain above the baseline, for how the incrementality of the marketing effects would have been in the treatment group if they had not been treated. In this case, it is the increase in conversion probability due to the treatment.

```
# Calculate the difference in probabilities for individuals actually treated with the baseline co
att <- mean(prob_treated[data$test == 1] - prob_cf[data$test == 1])

tau_lift <- att

cat("The incremental lift is", round(tau_lift*100, 2), "%\n")
```

```
The incremental lift is 0.74 %
```

The incremental lift is 0.74%, showing that by implying the treatment effect on the users, increase 0.74% probability on each individual to convert (to made a payment).

Profitability of the campaign

Get the true number of impressions by summing the number of impressions in the test group, for in the control group the impressions were indicated as those would have been assigned, not the real impressions have been executed.

```
# Define the average net contribution of 2000 NOK / purchase:
avg_net_contribution = 2000

# Summing the number of impressions in the test group:
no_impressions <- sum(data$impressions[data$test == 1])

# Calculate total cost according to number of impressions:
total_cost = (no_impressions/1000) * 50
```

```
# Calculate total revenue on total purchases:
revenue <- sum(data$purchase) * 2000
```

```
# Calculate the total profit:
profit <- revenue - total_cost

# Print the total profit:

cat('The total profit after the campaign: ', profit, 'NOK')
```

```
The total profit after the campaign:  1261602 NOK
```

The profit is approximately 1,261,601.5 NOK according to the observed data, shows that the campaign is profitable.

```
# Calculate profit from the control group:

profit_ctr_group = sum(data$purchase[data$test==0]) * 2000

# Profit from the test group:

profit_test_group = profit - profit_ctr_group

# Print the profit from the test group:
cat('The profit generated by test group:', profit_test_group, 'NOK')
```

```
The profit generated by test group: 1161602 NOK
```

The profit from the test group is 1,161,602 NOK indicates a good earning that was beneficial from the campaign and made a great contribution to the total profit.

The frequency effect of the ad exposure

For investigating the frequency effect of ad exposure in this case: Including $impressions$ variable in the logistic regression, as this variable indicates how many times the ad has been exposed to the individual, explains the general effect of ad exposure increase; as well as the interaction term of $test : impressions$ to see if there is a difference in treatment and control group towards ad exposure term.

```
# Added the 'impressions' variable and the interaction term 'test x impressions' into the regressi

model2 <- glm(purchase ~ test + impressions + test:impressions, data = data, family = binomial)

# Print the result:
tab_model(model2)
```

```
Profiled confidence intervals may take longer time to compute.
  Use `ci_method="wald"` for faster computation of CIs.
```

	purchase

Predictors	Odds Ratios	CI	*p*
(Intercept)	0.02	0.02 – 0.03	**<0.001**
test	1.15	0.84 – 1.60	0.399
impressions	1.04	1.02 – 1.06	**<0.001**
test × impressions	1.03	1.01 – 1.05	**0.010**
Observations	20000		
R^2 Tjur	0.050		

	purchase		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.02	0.02 – 0.03	**<0.001**
test	1.15	0.84 – 1.60	0.399
impressions	1.04	1.02 – 1.06	**<0.001**
test × impressions	1.03	1.01 – 1.05	**0.010**
Observations	20000		
R^2 Tjur	0.050		

The result table shows the effect of: treatment variable, number of impressions, and the interaction term of those two variables on the probability of making a purchase. Here is the interpretations given the table result:

The Intercept with the odds ratio of 0.02, shows that the odds of making a purchase of 2% given all other variables 0. Test and Impressions terms would be 0 when the individual is not in the treatment group, meaning the odds for purchasing in the control group is 2%. Its p-value is < 0.001 means that this intercept is statistically significant.

The 'test' term indicates the treatment assignment on individuals, with the odds ratio of 1.15 suggesting that being in the treatment group would increase the odds of purchasing by 15%, compared to being in the control group, holding other variables constant. p-value = 0.399 > 0.005 indicates a lack of statistical significance.

The term ‘impressions’ with an odds ratio of 1.04 suggests the odds of making a purchase would increase by 4%, holding other variables constant. p-value<0.001 means that this effect is highly significant. Demonstrating each ad impression added contributes to a 4% increasing in the odds of purchasing, emphasize its contribution in enhancing the likelihood of being converted.

The interaction term ‘test x impressions’ with an odds ratio of 1.03 indicates that the effect of impressions on the odds of making a purchase increases by 3% for the treatment group, compared to the control group. p-value=0.010 < 0.05 showing the significance of this variable, validates the hypothesis that the impact of the treatment on purchasing odds benefiting from the frequency of ad exposure. Moreover, the treatment makes incremental gains from each impression by increasing 3% of probability in making payment.

R-squared is 0.050, meaning that the explanatory power of this model is 5% in terms of the variance in the purchase outcome, as well as a weak predictive power.

While the treatment assignment alone seems to be significantly powerful in the previous model, with a high odds ratio (1.31), this model suggests a more balanced impact between different variables, of their contribution into the likelihood of a purchase, and the importance of considering ad exposure frequency (impressions) and its effect on the treatment group.

Conclusions

Opportunities & Obstacles

Obstacles:

- The provided data set contains the treatment variable and ad exposure frequency variable, which might not be enough to cover explanatory power, to make a more precise estimate on purchase decisions, as shown in the low R-squared value (5%). The external factors that were not captured in the data might influence the purchase decisions, and might indirectly introduce unobserved confounders.

- Estimates of impressions hypothetically for the control group might not be true, since this is not observable. Their behaviour learning based on this variable also might not be valid.

- The experiment was not enough to explain factors that influence the purchase decision from the control group. Because this analysis mainly focuses on the marketing campaign of showing advertisements, the conversion from the control group will serve as the baseline, but the effect that influences the purchase decision from the control group is still unobserved and there is room for further analysis if it was spillover effect from the treatment or not, or other exogenous factors.

- The imbalance of the treatment and control group assignment might affect the statistical power, cause skewness in the model distribution, and weaken in estimating the effects within these two groups.

Opportunities:

- Future experiments can benefit from this experiment, by taking into account more relevant factors that could affect the outcomes. Additional variables would enhance the performance of the experiment, in capturing the influence of different factors on customers in making the decisions. Ex: users' demographic, the loyalty card, ...

- There will also be room for better ad exposure in terms of users' behaviour learning. Some users were exposed more frequently but did not purchase, while there were users in the control group who saw no ad but purchased. Therefore a more advanced ad allocation for optimizing the resources should be considered. For instance: how long do the users spend to look at the ad, how to clarify the measurement of ad exposure that would have been shown to the control group,...

- A better allocation among treatment and control groups by re-evaluating the ratio would ensure more meaningful and valuable findings, statistically and economically significant, as well as representing the control group's characteristics more decently.

Potential Improvements

Besides the reallocation between treatment and control groups, the manager could consider observing the experiment for a longer period. It could provide more insights into long-term effects of ad exposure on making purchase decisions and customer loyalty.

Overall, this experiment has provided valuable insights into capturing the ad effectiveness and its great contribution to profitability, introducing exploring the potential for future experiments, helping to address the obstacles and possible improvements, the marketing manager could gain from them and making more informed decisions based on presented understandings of the campaign effectiveness.