

Business Context

A. Background

- Udacity is an educational organization offering massive open online courses.
- It has invested heavily in the course content, leaving the design team with sparse resources for quantitative data-driven decisions.
- The design team wants to introduce an add-on functionality on their subscription page, and they want to determine if the decision can be supported by quantitative studies
- The new feature displays a precaution message to users who enrol for the free trial option- this will only be displayed to users who do not meet the criteria of allocating five hours per week

B. Intuition behind the new feature design

- The design team's hypothesis is that a new feature will set clearer expectations for students upfront, thus reducing the number of frustrated students who left the free trial because they did not have sufficient time. This should be accompanied by a non-substantial reduction in the number of students who continue past the free trial and eventually complete the course
- If this hypothesis holds true, it would lead to a more positive experience for users, and also help Udacity in capacity planning of coaches, with a more realistic figure for students who intend to complete the course (visual walkthrough and detailed explanation given in next section)

C. Causal Question

The design team wants to assess if a proposed new feature improves user experience and elevates retention. This will be measured through the maximization of two metrics - course completion and retention of students post free trial period.

D. Importance

It is useful to understand the causal impact of this design feature, because this design feature is intended to increase positive customer experience and bring in more revenue per user who enrolls for a free trial. Elevating these figures will consequently translate to more revenue for the Udacity.

Design of Experiment

A. Website user experience before the experiment

Visitors who experienced this flow are a part of Control Group

<IMAGE: Initial Flow>

1. Udacity courses currently have two options on the course overview page: "Access Course Materials" and "Start Free Trial"
2. If a student clicks on **Start Free Trial**, they will be asked to enter their credit card details, and subsequently enrolled in a free trial of the paid version of the course. After 14 days, they will be charged automatically, unless they cancel the subscription first.
3. If the student clicks **Access Course Materials**, they will be able to access course materials (like viewing videos or taking the quiz for lessons) for free, but they will not receive a verified certificate or coaching support. Additionally, they cannot submit their final project for personalised feedback.

B. Website user experience change, after the experiment

Visitors who experienced this flow are a part of Test or Treatment Group

<IMAGE: Final Flow>

- In this experiment, the enrolment process remains the same for users who click *Access Course Materials*. While one extra enrolment step is added for users who click *Start Free Trial*, as treatment of this experiment.
- The treatment is a pop-up window that asks users to enter the number of hours they are likely to commit to course learning in Udacity, during the free trial.
- If the user inputs a number larger than five, then he or she will be sent back to the usual enrolment process
- If the user inputs a number smaller than five, then a message will pop up, informing the user that Udacity generally requires more commitment to the study and suggesting him or her to choose *Access Course Materials* option instead.

C. Platform being used to conduct the experiment

- The Udacity design team is implementing this design change through Facebook Atlas platform
- Atlas follows a black-box method to assign randomized test and control groups, which are shown two different visual layouts or

website user experience

- Atlas provides the required data, for measuring the effectiveness of the experiment
- Currently, Udacity is a non-premium member and hence does not have full-blown access to the entire clickstream data captured during the experiment - both in terms of coverage as well as granularity
- The unit of diversion is a visitor cookie, although if the student enrolls in the free trial, they are tracked by user-id from that point onwards. The same user-id can only enrol for a free trial once. For users that do not enrol, their user-id is not tracked in the experiment, even if they were signed in when they visited the course overview page.

D. Dataset Overview

Unit of observation: The data is on a daily level. Each row of record has aggregated each day's pageviews, clicks, enrolments and payments.

Time period over which the data is observed: Data set contains data from Oct 11th to Nov 2nd - 23 days in total

E. Important features in the dataset

Considering the limited budget, the Facebook Atlas platform only gives visibility to five attributes, for measuring the efficacy of our experiment

- **Date:** Date on which the data is recorded
- **Page Views:** Number of unique users to view the course overview page that day
- **Clicks:** Number of unique users to click the course overview page that day
- **Enrolments:** Number of user-ids to enrol in the free trial that day
- **Payments:** Number of user-ids who enrolled on that day to remain enrolled for 14 days and thus make a payment

In addition to the above, we are also creating derived variables for Weekday/Weekend flag based on Date variable - **Weekdays:** Binary variable, 1 indicates weekdays - **Weekend:** Binary variable, 1 indicates weekend

Treatment variable(X) : Binary variable, 0 indicates control group, 1 indicates treatment group

F. Metrics for measuring Outcome

We are considering three evaluation measures to get a holistic read on the effects of the experiment, and a causal impact of the design feature:

Measure	Description	Calculation
Gross Conversion	Number of enrolled users of the free trial divided by the number of cookies clicking on free trial button	$Enrols / Clicks$
Net Conversion	Number of paying users divided by the number of cookies clicking on free trial button	$Payments / Clicks$
Retention	Number of paying users divided by the number of enrolled users of the free trial	$Payments / Enrols$

Threats to Causal Inference

A. Interference

Even though the experiments are randomized, the users who are a part of the experiment might know each other. The behaviour of people towards the experiments might deviate from natural free choice in that case

B. Measurement Error

Some users may pay for the courses after free trial if they forget to cancel the enrolment. This might lead to a measurement error of one of our independent variables, which is enrolment.

Analysis of Experiment

We wanted to see if our experimentation setting is ideal enough to interpret the results.

A. Validating assumptions for ideal experimentation setting

Randomization Check:

- **Why:** Since the randomization task was performed by a black-box model in Facebook Atlas platform, we wanted to double-check if the treatment and control population were truly randomized for the experiment
- **Methodology:** First, we checked if our test and control samples were similar enough. This was done to ensure that nothing other than treatment was different in both samples so that the exact effect of the treatment can be observed. The results suggested appropriate randomization in test and control.

<IMAGE: Daily Page Views>

<IMAGE: Daily User Clicks>

<IMAGE: Daily Click Through Rate>

Comment: We have successfully performed the sanity checks, to assess if the test and control groups are not inherently biased in terms of randomization

Sample size:

Using this test, we wanted to assess if our sample size was sufficient enough to run our experiment and conclude with its inferences. This suggested that due to insufficient sample size, we could not interpret the results as per our findings.

B. Results and findings

We ran a linear regression for the three evaluation metrics respectively to study the casual relationship between the metrics and pageviews, clicks, weekday/weekend flag, and the treatment. Following are our findings:

<IMAGE: Results for statistical significance on outcome variables.png>

<IMAGE: Gross Conversion.png>

Results

- For gross conversion, we found that the treatment group had 0.019 decrease in gross conversion when compared with the control group.
- For net conversion, the new feature had no significant impact.
- For retention, similar to net conversion, the new feature had no significant impact on it.

Managerial Interpretation:

- A statistically and practically significant decrease in Gross Conversion was observed, in the group which was showed a warning message, but with no significant differences in Net Conversion.
- This translates to a decrease in enrollment not coupled to an increase in students staying for the requisite 14 days to trigger payment

Limitations:

A. Duration of the experiment:

The length of the available data is one month. We have 37 rows of records for both control and test groups, it may be insufficient to establish a measure of success (causal relationship) with strong certainty.

B. Omitted Variable(s):

There might be other unaccounted variables which influence the three performance evaluation metrics, for example, inequitable exposure of test and control populations to some advertising campaigns run by Udacity.

C. Ethical Consideration:

By showing a calculated warning to a section of users in the test group, and not to the control group, we are revealing more information to them in order to make a decision. This is putting the users in the control group at a disadvantage. Informed consent has also not been taken from users.

D. Holistic Consideration:

We are missing customer attribute information. As such, even if the treatment yields an overall positive effect, we do not know if it has a negative outcome on a certain section of users. In other words, there is no way to breakdown the analysis by customer segments and observe the variation in each segment

Recommendations:

A. Increase investment and resources into conducting experiments

The results of the power test showed how severely our dataset is underpowered to study the causal effects of the new design feature on the Udacity website. Our recommendation is to become a Premium user of Facebook Atlas platform, so that the design team can increase both the duration of the experiment as well as the granularity and coverage of data - With access to user-level data, it would be possible to study the effects of the experiment on segments of the users. - Additionally, it would allow us to conduct deep-dive experiments into the how the course completion rate and retention rate differed for users who belonged to the two categories of ≥ 5 hours commitment per week and < 5 hours commitment per week

B. In case of persisting budget constraints

A statistically significant decrease in Gross Conversion was observed but this was not accompanied with any significant differences in Net Conversion. This translates to a decrease in enrolment is not coupled with an increase in students staying for the requisite 14 days to trigger payment. Considering this observation, our recommendation is not to launch the feature, but rather to pursue other experiments. - This could be focused on the user journey after enrolment, and before scheduled payment, to observe the user engagement through the 14 day trial period. This may require intervention measures like personalised coaching and feedback, reminder to devote requisite hours per day, or pairing users and forming groups to increase motivation for completion - This could also focus on making the warning message more personalized, with a pre-requisite list if the user is particular about exploring an individual course

Appendix:

```
library(ggplot2)
library(tidyr)
library(lubridate)

treat = read.csv('experiment_data.csv')
control = read.csv('control_data.csv')

# metrics

# set baseline value for each metrics - how these metrics behave before the change
# Udacity gives the following rough estimates for these metrics - traffic on a daily level, population
pageviews = 40000 # -> invariate, Number of unique cookies to view the course overview page that day.
clicks = 3200 # -> invariate, Number of unique cookies to click the course overview page that day.
enrolls = 660 # Number of user-ids to enroll in the free trial that day.
click_through_rate = 0.08 # clicks / pageviews -> invariate
Gross_Conversion = 0.20625 # enrolls / clicks -> evaluate, Dmin = 0.01
Retention = 0.53 # payments / enrolls -> evaluate, Dmin = 0.01
Net_Conversion = 0.109313 # payments / clicks -> evaluate, Dmin = 0.0075

## data preparation
control1 = control[1:23,]
treat1 = treat[1:23,]
control1$treatment = 0
treat1$treatment = 1
data = rbind(treat1, control1)
data = separate(data, Date, c('Week', 'Date'), sep = ',')
data$click_through_rate = data$Clicks / data$Pageviews
data$Gross_Conversion = data$Enrollments / data$Clicks
data$Retention = data$Payments / data$Enrollments
data$Net_Conversion = data$Payments / data$Clicks
data$treatment = as.factor(data$treatment)
# create dummy variable for weekday & weekends, set the interaction of treatment and weekday/weekend
data$Date = parse_date_time2(data$Date, orders = "md")
data$weekdays = 0
data[data$Week %in% c('Mon', 'Tue', 'Wed', 'Thu', 'Fri'),]$weekdays = 1
data$weekend = 0
data[data$Week %in% c('Sat', 'Sun'),]$weekend = 1

# How big a sample we would need to detect the practical change they hope to find?
## Gross conversion rate baseline = 0.20625 and least detectable change = 0.01 -> delta = 0.0020625
delta_grossc = 0.20625 * 0.01
power.t.test(n=NULL, type=c("two.sample"), power=0.8, sig.level=0.1, delta=delta_grossc, sd=sd(data$Gross_Conversion))
```

```

# 6277

## retention rate baseline = 0.53 and least detectable change = 0.01 -> delta = 0.0053
delta_retention = 0.53 * 0.01
power.t.test(n=NULL,type=c("two.sample"),power=0.8,sig.level=0.1,delta=delta_retention, sd=sd(data$Retention
))
# 5539

## Net conversion rate baseline = 0.109313 and least detectable change = 0.0075 -> delta =
delta_netc = 0.109313 * 0.0075
power.t.test(n=NULL,type=c("two.sample"),power=0.8,sig.level=0.1,delta=delta_netc, sd=sd(data$Net_Conversion
))
# 17201

# current sum of pageviews for treatment and control group
treatment_sumviews = sum(data[data$treatment==1,]$Pageviews)
control_sumviews = sum(data[data$treatment==0,]$Pageviews)
total_views = treatment_sumviews + control_sumviews
total_views

# descriptive analysis
ggplot(data, aes(Date, Pageviews, col = treatment)) +
  geom_line()+
  xlab('Date') +
  ylab('PageViews') +
  ggtitle('Pageviews for Control and Test Group')+
  theme(plot.title = element_text(size=22),
        axis.title.x = element_text(size=14),
        axis.title.y = element_text(size=14),
        legend.title = element_text(size=14))+
  theme(plot.title = element_text(hjust = 0.5))

ggplot(data, aes(Date, Clicks, col = treatment)) +
  geom_line()+
  xlab('Date') +
  ylab('Clicks') +
  ggtitle('Clicks for Control and Test Group')+
  theme(plot.title = element_text(size=22),
        axis.title.x = element_text(size=14),
        axis.title.y = element_text(size=14),
        legend.title = element_text(size=14))+
  theme(plot.title = element_text(hjust = 0.5))

ggplot(data, aes(Date, click_through_rate, col = treatment)) +
  geom_line()+
  xlab('Date') +
  ylab('Click Through Rate') +
  ggtitle('Click Through Rate for Control and Test Group')+
  theme(plot.title = element_text(size=22),
        axis.title.x = element_text(size=14),
        axis.title.y = element_text(size=14),
        legend.title = element_text(size=14))+
  theme(plot.title = element_text(hjust = 0.5))

## Randomization check
t.test(Pageviews ~ treatment, data = data) # p-value = 0.887
t.test(Clicks ~ treatment, data = data) # p-value = 0.9477
t.test(click_through_rate ~ treatment, data = data) # p-value = 0.8753

## Testing
# Gross Conversion
m1 = lm(Gross_Conversion ~ treatment + Pageviews + Clicks + weekend, data = data)
summary(m1)

# Net Conversion
m2 = lm(Gross_Conversion ~ treatment + Pageviews + Clicks + weekend, data = data)
summary(m2)

# Retention
m3 = lm(Gross_Conversion ~ treatment + Pageviews + Clicks + weekend, data = data)
summary(m3)

```

```
# GC - interactions
m4 = lm(Gross_Conversion ~ treatment + treatment * Pageviews + treatment * Clicks, data = data)
summary(m4)

# Net Conversion - interactions
m5 = lm(Net_Conversion ~ treatment + treatment * Pageviews + treatment * Clicks, data = data)
summary(m5)

# Retention - interactions
m6 = lm(Retention ~ treatment + treatment * Pageviews + treatment * Clicks, data = data)
summary(m6)
```

Loading [MathJax]/jax/output/HTML-CSS/jax.js