

A/B test Final Project

Dominic Del Rosario

Experiment Design

Metric Choice

Number of cookies: This is an invariant metric because this number does not depend on the change Udacity wants to test. The number of unique cookies to visit the course overview won't change.

Number of user-ids: This metric won't be an invariant metric because this could depend on the change Udacity wants to test. Also, it won't be an evaluation metric because the number by itself is not useful as a ratio of this number to cookies, clicks, etc because we would want affect an even distribution in the experiment and control groups. A raw number would not necessarily do this.

Number of clicks: This is an invariant metric because this number does not depend on the change Udacity wants to test. The number of unique cookies to click "Start free trial" occurs before Udacity's free trial screener.

Click-through probability: This is an invariant metric because this number does not depend on the change Udacity wants to test. Additionally, this ratio is the above invariant metrics I just mentioned (number of cookies and number of clicks).

Gross conversion: This is an evaluation metric because this number depends on the change Udacity wants to test. This conversion is a metric Udacity wants to know. They want to know who continues to the free trial after the free trial screener.

Retention: This is an evaluation metric because this number depends on the change Udacity wants to test. Secondly, this conversion is a metric Udacity wants to know. They want to know who continues past free trial period and make a payment.

Net conversion: This is an evaluation metric because this number depends on the change Udacity wants to test. Finally, this conversion is a metric Udacity wants to know. They want to know who continues past free trial period and make a payment.

To launch the experiment, I will be looking for the following:

Gross conversion - can decrease because we are looking to decrease the number of unprepared students.

Retention - we just don't want this metric to significantly decrease.

Net conversion - we just don't want this metric to significantly decrease.

Measuring Standard Deviation

Gross conversion: 0.0202

Retention: 0.0549

Net conversion: 0.0156

I think the analytic estimates would be comparable to the empirical variability in two of the three cases. In the case of gross conversion and net conversion, we are measuring the number of user-ids to some denominator (number of cookies to click "start free trial"), and the unit of diversion is cookies, so the estimates should be comparable. In the case of retention, the unit of analysis is user-id and the unit of diversion is a cookie, the estimates may not be comparable.

Sizing

Number of Samples vs. Power

I will not be using a Bonferroni correction during the analysis phase. We will be looking for a significant change in all of the metrics to launch a change.

Pageviews: 685,275

Duration vs. Exposure

If diverted 50% of the traffic to this experiment which would take 35 days to run. One reason is for safety in that we want an even number to run through the experiment. However, the client expects the time frame to be a few weeks. I would divert 60% of the traffic which would take 29 days to run. Technically, that falls within "a few weeks" range. If we divert a lower percentage than 50%, we will need to run the experiment much longer than 35 days.

There is a risk to the free trial screener but it should be low. Since there is no harm to the user and we are not collecting sensitive information, such as medical information, this should be low risk.

During the computation of number of pageviews and days to run experiment, it was deemed that retention wouldn't be a good metric to use. It would require 4.7 million pageviews and 237 days to run experiment which would be too long to run.

Experiment Analysis

Sanity Checks

	Con Interval -	Con Interval +	Observed	Passes
Cookies	0.5012	0.5006	0.5006	yes
Clicks	0.4959	0.5041	0.5005	yes
Click-through probability	-0.0013	0.0013	0.0001	yes

Result Analysis

Effect Size Tests

	Con Interval -	Con Interval +	Statistical Sig	Practical Sig
Gross conversion	-0.0291	-0.0120	yes	yes
Net Conversion	-0.0116	0.0019	no	no

Sign Tests

	Sign test p-value	Statistically sig
Gross conversion	0.0026	yes
Net conversion	0.6776	no

Summary

As stated above, I will not be using a Bonferroni correction. We will be looking for a significant change in all of the metrics to launch a change.

Recommendation

Regarding the evaluation metric, gross conversion, at a 95% confidence interval, we found the difference between the experiment and control groups to be statistically and practically significant. For the evaluation metric, net conversion, at a 95% confidence interval, we did not find the difference between the experiment and control groups to be statistically and practically significant. The confidence interval of net conversion is (-0.0116, 0.0019). Since it includes 0,

it is not statistically significant. Since $-d_{min}$ is -0.0075 and lies within the confidence interval, there is evidence that net conversion may decrease.

For that reason, I would recommend to dig deeper into the experiment. We do see a significant difference for gross conversion of number of cookies to click "Start free trial" to completing checkout and enrolling in the free trial. The decrease in gross conversion is aligned with the goals. However, Udacity also wants to know that the Net conversion is significant as well. They want to know that the overall student experience has been improved and expectations are set up-front correctly. Since there is a chance net conversion could decrease, it causes for further investigation.

Follow-Up Experiment

Lengthen the trial period to 21 days

An experiment that I would try is to lengthen the trial period to 21 days from 14 days. It has been said that it takes 21 days to make or create a habit. That habit can range from exercising, a diet change, waking up early or even studying. A lot of companies know this and that's why you will see companies like Netflix, Hulu, Spotify offer free one-month trials. They know a certain percentage will enjoy the service and value it after 21 days to a month and continue and pay for the service.

The hypothesis is that this allows the students to access the course materials and even go through the process of finishing a project and getting it reviewed. That is critical because I had no idea I could even do this course after my 7 day trial.

The unit of diversion is a cookie, but if they enroll in free trial, they are tracked by user-id from that point forward. If they don't enroll, their user-id is not tracked in the experiment.

One invariant metric would be number of cookies or the number of unique cookies to view the course overview page. This is an invariant metric because this number does not depend on the change Udacity wants to test. The number of unique cookies to visit the course overview won't change.

I would be tracking gross conversion and net conversion just like with the free trial screener. I want to know who is continuing on to the free trial but I also want to know who is making at least a payment.