

Experiment Design

Metric Choice

Firstly, I would say the purpose of the experiment is that there would be more effectiveness for the quality of education by reducing a number of students who would quit while taking free trial.

Invariant metric:

Number of cookies: Chosen. In the course overview page, it would not be affected by the pop up message.

Number of clicks on "Start free trial": Chosen. It is also the process before the alert message. So, it would not change for a control group and an experiment group

Click-through-probability on "Start free trial": Chosen. This is the number of click divided by number of cookies, which is not affected by the change.

Number of user id: Not chosen. It is affected by the change. The number of user id would be changed due to the pop up message

Gross conversion: Not chosen. Numerator, which is user id who completed check out, would be changed for a control group and experiment group.

Net conversion: Not chosen. Same as Gross conversion. Numerator would be changed by the message.

Retention: Not chosen. Same as above. Numerator changed by the message.

Evaluation metric:

Number of cookies: invariant metric

Number of clicks on "Start free trial": invariant metric

Click-through-probability: invariant metric

Number of user id: It could be an evaluation metric but not chosen because it is not normalized.

It doesn't consider how many visited website. Number of user ID could check whether the number increased or not before and after the change.

Retention: It couldn't measure whether the number of user increased or not so not chosen. A number of user id who paid divided by a number of user id who enrolled. No common with an invariant metric. Same as number of user id affected by people who visited website some special day.

Gross conversion: Chosen. It would show whether a number of student who enrolled the free trial increased or not by the change. We expect that the gross conversion would be decrease due to the message. Students who don't have enough time would not register.

Net conversion: Chosen. It would show whether a number of student who made their first payment increased or not. To launch the experiment, it is expected that the net conversion would not be decrease.

For each metric, explain both why you did or did not use it as an invariant metric and why you did or did not use it as an evaluation metric. Also, state what results you will look for in your evaluation metrics in order to launch the experiment.

Measuring Standard Deviation

Calculated standard deviation:

Gross conversion = 0.0202

Net conversion = 0.0156

Analytic estimate comparable to empirical variability

List the standard deviation of each of your evaluation metrics. (These should be the answers from the "Calculating standard deviation" quiz.)

Gross conversion = Enrolled the class/ unique users clicked "Free trial"

Unit of diversion: cookie that clicked "Free trial"

Unit of analysis: cookie

Net conversion = Made a first payment / unique users clicked "Free trial"

Unit of diversion: cookie that clicked "Free trial"

Unit of analysis: cookie

Both evaluation metrics have the same unit, which is cookie. Analytic estimate is comparable to the empirical variability in both cases.

Sizing

Number of Samples vs. Power

No. Not used Bonferroni.

This experiment would be launched in case both evaluation metric changed as expected. Gross conversion should be decrease and Net conversion should be maintained. As these two evaluation met our expectation, then it would be worth to launch it. So, we don't need to use Bonferroni to reduce a risk. If we could launch any of our metric met what we expected, then we need to use Bonferroni to reduce a risk

685325 number of page view needed

Duration vs. Exposure

I decided to use 1 fraction of traffic for the experiment.

Then, length of days needed 18 days. Almost a half month.

It could be the shortest length of days for the test but it couldn't get as many information as needed because it took 14days to monitor the change for an anonymous person to make their first payment.

Indicate what fraction of traffic you would divert to this experiment and, given this, how many days you would need to run the experiment. (These should be the answers from the "Choosing Duration and Exposure" quiz.)

There is no risk for the experiment. In the point of sensitivity, it doesn't include any political, personal history, or sexual issues. It just asked how much time students could devote for the class. Also, it doesn't anyone to get hurt during the free trial. It doesn't force students to study more than 5 hours something like this, which might result in someone hurt. The experiment is totally okay with a risk.

Give your reasoning for the fraction you chose to divert. How risky do you think this experiment would be for Udacity?

Experiment Analysis

Sanity Checks

For each of your invariant metrics, give the 95% confidence interval for the value you expect to observe, the actual observed value, and whether the metric passes your sanity check. (These should be the answers from the "Sanity Checks" quiz.)

Number of cookies → LB: 0.4988, UB: 0.5012, Observed: 0.5006, Passed

Number of clicks on start free trial → LB: 0.4959, UB: 0.5041, Observed: 0.5005, Passed

Passed all

For any sanity check that did not pass, explain your best guess as to what went wrong based on the day-by-day data. **Do not proceed to the rest of the analysis unless all sanity checks pass.**

Result Analysis

Effect Size Tests

Gross conversion → LB: -0.0291, UB: -0.0120, statistical, practical significance

Net conversion → LB: -0.0116, UB: 0.0019, None of them

For each of your evaluation metrics, give a 95% confidence interval around the difference between the experiment and control groups. Indicate whether each metric is statistically and practically significant. (These should be the answers from the "Effect Size Tests" quiz.)

Sign Tests

Gross conversion → P-value: 0.0026, statistical significance

Net conversion → P-value: 0.6776, None

For each of your evaluation metrics, do a sign test using the day-by-day data, and report the p-value of the sign test and whether the result is statistically significant. (These should be the answers from the "Sign Tests" quiz.)

No. Not used Bonferroni.

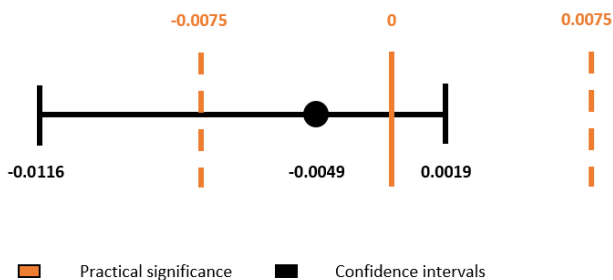
This experiment would be launched in case both evaluation metric changed as expected. Gross conversion should be decrease and Net conversion should be maintained. As these two evaluation met our expectation, then it would be worth to launch it. So, we don't need to use Bonferroni to reduce a risk. If we could launch any of our metric met what we expected, then we need to use Bonferroni to reduce a risk

Recommendation

Make a recommendation and briefly describe your reasoning.

Gross conversion behaved as expected. The number of students who enrolled the free trial decreased. But Net conversion is hard to say that met the expectation because the number of students who paid their class might be decreased in some cases.

NET CONVERSION



We expected gross conversion would decrease and net conversion maintained. Net conversion could be below the negative of the confidential intervals. So, it couldn't meet our expectation at least net conversion needed to be maintained.

Not launch the change yet. Needed more test for net conversion.

Net conversion needed more test I think. The result for the net conversion presented not that much affected by the change.

Follow-Up Experiment

Give a high-level description of the follow up experiment you would run, what your hypothesis would be, what metrics you would want to measure, what your unit of diversion would be, and your reasoning for these choices.

Follow-Up Experiment: How to Reduce Early Cancellations: If you wanted to reduce the number of frustrated students who cancel early in the course, what experiment would you try?"

Early cancellation. Why.

1. Not fun.
2. No time.
3. No achievement
4. Cost effectiveness
5. Too difficult.

Okay. Here is a thing.

Reduce the number of cancellations. The best is that giving them money back bonus way more than it is now, especially in the beginning of the course. I will take 50% of the money back when it completed within a year and there is like a bonus \$25 money back when I completed each project. Same like this way, give them some discount card when they completed each lecture! Not the project!. Give them a reward in the beginning something like \$5 or \$10 saving coupon in their account and apply them when they make their first payment.

So, Let's give them a reward whenever they completed each classes to their account not usable unless they make their first payment.

Unit of diversion: user id

Hypothesis: the number of user id quitting their trial in the beginning would be decreased by giving them a reward.

metric:

1. Invariant: A number of user id who enrolled the free trial
2. A number of user id who remained for one week
3. Evaluation: A number of user id who remained for one week divided by a number of user id who enrolled the free trial

The evaluation metric expected to be increased