P7 A/B TESTING

Experiment Overview: Free Trial Screener

At the time of this experiment, Udacity courses currently have two options on the home page: "start free trial", and "access course materials". If the student clicks "start free trial", they will be asked to enter their credit card information, and then they will be enrolled in a free trial for the paid version of the course. After 14 days, they will automatically be charged unless they cancel first. If the student clicks "access course materials", they will be able to view the videos and take the quizzes for free, but they will not receive coaching support or a verified certificate, and they will not submit their final project for feedback.

In the experiment, Udacity tested a change where if the student clicked "start free trial", they were asked how much time they had available to devote to the course. If the student indicated 5 or more hours per week, they would be taken through the checkout process as usual. If they indicated fewer than 5 hours per week, a message would appear indicating that Udacity courses usually require a greater time commitment for successful completion, and suggesting that the student might like to access the course materials for free. At this point, the student would have the option to continue enrolling in the free trial, or access the course materials for free instead.This screenshot shows what the experiment looks like.

The hypothesis was that this might set clearer expectations for students upfront, thus reducing the number of frustrated students who left the free trial because they didn't have enough time—without significantly reducing the number of students to continue past the free trial and eventually complete the course. If this hypothesis held true, Udacity could improve the overall student experience and improve coaches' capacity to support students who are likely to complete the course.

The unit of diversion is a cookie, although if the student enrolls in the free trial, they are tracked by user-id from that point forward. The same user-id cannot enroll in the free trial twice. For users that do not enroll, their user-id is not tracked in the experiment, even if they were signed in when they visited the course overview page.

Experiment Design

Metric Choice

List which metrics you will use as invariant metrics and evaluation metrics here. (These should be the same metrics you chose in the "Choosing Invariant Metrics" and "Choosing Evaluation Metrics" quizzes.)

Invariant Metrics :

- Number of cookies
- Number of clicks

Evaluation Metrics :

- Gross conversion
- Retention
- Net conversion

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.

INVARIANT METRICS

Number of cookies:
That is, number of unique cookies to view the course overview page. This is page comes before the experiment starts and independent from the experiment . This is evenly distributed among the experiment and control group .

Number of clicks:
That is the number of unique cookies to click the "Start free trial" button which happens before the free trial screener is trigger so it is independent from the experiment and It is evenly distributed among the Control and Experiment groups.

EVALUATION METRICS

Gross conversion:
That is, number of user-ids to complete checkout and enroll in the free trial divided by number of unique cookies to click the "Start free trial" . This is the direct result of the experiment thus suitable to be an evaluation metric .So for our experiment we will test whether adding the time commitment option will make any change in the gross conversion of the experiment group from the control group and whether the change is significant .

Retention:
That is, number of user-ids to remain enrolled past the 14-day boundary and thus make at least one payment divided by number of user-ids to complete checkout. This is also suitable to be an evaluation metric . Retention is highly dependent on the experiment . High retention will mean higher number of enrolled students who have paid at least once . So for our experiment we will test whether adding the time commitment option will make any change in the Retention of the experiment group from the control group and whether the change is significant .

Net conversion:
That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by the number of unique cookies to click the "Start free trial" button. Net Conversion will show the number of students who will be paid students after the 14 day free trial . So for our experiment we will test whether adding the time commitment option will make any change in the net conversion of the experiment group from the control group and whether the change is significant .

NOT TAKEN AS EVALUATION OR INVARIANT METRICS

Click-through-probability:
That is, number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page.This won't be dependent on the experiment and it is also not guaranteed that the click through probability will be distributed evenly among the control and experiment group .

Number of user-ids:
That is, number of users who enroll in the free trial . This is not an ideal metric for both evaluation and invariant . This will be dependent on the experiment but can vary drastically between the control and experiment group .

Measuring Standard Deviation

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

METRICS NAME	STANDARD DEVIATION
Gross Conversion	0.0202
Retention	0.0549
Net Conversion	0.0156

For each of your evaluation metrics, indicate whether you think the analytic estimate would be comparable to the the empirical variability, or whether you expect them to be different (in which

case it might be worth doing an empirical estimate if there is time). Briefly give your reasoning in each case.

Both Gross and Net Conversion are measured per number of cookies clicked in the 'Start Free Trial' button, here number of cookies is also the unit of diversion . Here, the unit of diversion is equal to unit of analysis . So the analytical estimate would be comparable to the empirical variability.

Here Retention is measured per the number of user-ids . Here, the unit of diversion is not equal to unit of analysis . So the analytical estimate and the empirical variability will be different.

Sizing

Number of Samples vs. Power

Indicate whether you will use the Bonferroni correction during your analysis phase, and give the number of pageviews you will need to power you experiment appropriately. (These should be the answers from the "Calculating Number of Pageviews" quiz.)

No I will not use the Bonferroni correction during the analysis phase because the metrics in the experiment has high correlation .

Metric Name	Dmin	Baseline Conversion	Samples Needed	Total PageViews
Gross Conversion	0.01	20.625 %	25835	645875
Retention	0.01	53 %	39115	4741212
Net Conversion	0.0075	10.93125 %	27413	685325

Duration vs. Exposure

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.)

Percentage of Traffic	50%
Days Required	35

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

Here I chosed 50% traffic for this experiment . Already registered students won't be affected by this experiment . Only the new users will be affected so it won't be High Risk 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:

Pageviews in Control Group	345543
Pageviews in Experiment Group	344660
Total Pageviews(N)	690203
Probability of Pageview to be in Control Group or Experiment Group	.5
Standard Error(SE)	sqrt(.5*(1-.5)/N) = 0.0006018
Margin of Error(m)	SE*Z = SE*1.96 = 0.0011796
Confidence Interval(CI)	0.5 (+-) m = [0.4988,0.5012]
Observed Value	345543/690203 = 0.5006

Observed Value 0.5006 in CI [0.4988,0.5012]
Hence Number of Cookies PASS Sanity Check

Number of Clicks:

Clicks in Control Group	28378
Clicks in Experiment Group	28325
Total Pageviews(N)	56703

Probability of Pageview to be in Control Group or Experiment Group	.5
Standard Error(SE)	sqrt(.5*(1-.5)/N) = 0.0021
Margin of Error(m)	SE*Z = SE*1.96 = 0.0041
Confidence Interval(CI)	0.5 (+-) m = [0.4959,0.5041]
Observed Value	28378/56703=0.50046

Observed Value 0.50046 in CI [0.4959,0.5041]
Hence Number of Clicks PASS Sanity Check

Result Analysis

Effect Size Tests

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.)

GROSS CONVERSION:

Clicks Control Group (N_{cont})	**17293**
Enrollment Control Group (X_{cont})	**3785**
Clicks Experiment Group(N_{exp})	**17260**
Enrollment Experiment Group(X_{exp})	**3423**
Gross Conversion Control Group(P_{cont})	**0.218875**
Gross Conversion Experiment Group(P_{exp})	**0.19832**
D-Hat	**-0.02055**
Pooled Probability (P_{pool})	**0.2086**
Standard Error (SE)	**0.00437**
Margin of Error (m)	**0.00857**
Confidence Interval	**[-0.02912,-0.01199]**

Confidence Interval **[-0.02912,-0.01199]**

Statistically Significant - Dont Have 0 in CI
Practically Significant - CI does not contain d_{min}

NET CONVERSION:

Clicks Control Group (N_{cont})	**17293**
Payment Control Group (X_{cont})	**2033**
Clicks Experiment Group(N_{exp})	**17260**
Payment Experiment Group(X_{exp})	**1945**
Net Conversion Control Group(P_{cont})	**0.11756**
Net Conversion Experiment Group(P_{exp})	**0.11268**
D-Hat	**-0.0049**
Pooled Probability (P_{pool})	**0.1151**
Standard Error (SE)	**0.00343**
Margin of Error (m)	**0.0067**
Confidence Interval	**[-0.0116 , 0.0018]**

Confidence Interval [-0.0116 , 0.0018]
Not Statistically Significant - CI contains 0
Not Practically Significant - CI contains d_{min}

Sign Tests

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.)

GROSS CONVERSION

Number of Success	19
Number of Experiments	23
Probability	0.5
Two Tailed P-value	0.0026

As P-value (0.0026) < Alpha Level (0.025)
Statistically and practically significant.

NET CONVERSION

Number of Success	13
Number of Experiments	23
Probability	0.5
Two Tailed P-value	0.6776

As P-value (0.6776) > Alpha Level (0.025)
Statistically and practically Insignificant.

Summary

State whether you used the Bonferroni correction, and explain why or why not. If there are any discrepancies between the effect size hypothesis tests and the sign tests, describe the discrepancy and why you think it arose.

No I will not use the Bonferroni correction because the metrics in the experiment has high correlation .We will only launch if we get a significant amount of change over all metrics . In this case there is no need of Bonferroni Correction.

Recommendation

Make a recommendation and briefly describe your reasoning.

As we can see that Gross Conversion change is both statistically and practically significant and there is a high chance that the Gross Conversion will go down . By launching this feature we will discourage those students who cannot commit a certain amount of time every week to this course so our overall enrollment will go down which include the revenue too , though there is no significant change of Net Conversion both statistically and practically . YES Udacity should launch this feature , which will decrease the number of enrollments who don't have enough time for this course , so more resource could be allocated for the students who can . Adding this feature won't be a high risk in revenue because we don't see any significant change in Net Conversion both Statistically and Practically .

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.

In the follow up experiment I would like that to have a quiz for the students for the First Free trail week that checks the knowledge a student should have before starting the Data Analyst Nanodegree which includes basic python knowledge , working in the Ipython notebook etc.

I will evenly distribute the quiz between the control and the experiment groups , the students in the control group will not have the quiz on the other hand the experiment group students will have this quiz .

Null Hypothesis (H_0) :
Adding the quiz don't increase Retention .

Metrics:
INVARIANT
Number of Unique users enrolled in the Course as Free Trial
As the number of unique users enrolled in the course is not dependent on the experiment as the experiment is staged after the user Starts the free Trial.

EVALUATION
Number of unique users that actually makes the first payment .
As the number of unique users who make the first payment which will be done after the Trial Period where the experiment is staged . So these are chosen as Evaluation Metrics .

Unit of Diversion:
User IDs will be the unit of diversion because here we will only experiment on the users who actually enrolled for the course as a Free Trial .