

## Template Format

This template can be used to organize your answers to the final project. Items that should be copied from your answers to the quizzes should be given in [blue](#).

## 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 index](#): The number of Cookie, clicks, click-through probability

[Evaluation](#): The total conversion, 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.

Cookie Number: It is good invariant metrics because the users see it before the experiment, and therefore it is independent.

User ID Number: This is not a good indicator for evaluation or invariant. Not a good invariant because it is dependent with the experiment. on the number of registered users. Not an ideal evaluation because the number of visitors may be different in the experimental and control groups, which likely to skew the results.

Number of clicks: Good invariant metrics, because the user clicks before the experiment, and therefore independent.

CTP: good invariant metrics because users clicks before the experiment, and therefore independent.

Gross conversion: This is not a good invariant indicator, because user enroll the free trial experiments dependent on experiment. Good evaluation metrics, because it is directly dependent on the effect of the experiment, and allow us to show whether we are successful in reducing the cost of enrollments less likely to become paying customers.it can be used as an

evaluation metric to check if the experiement makes a significant difference in the enrolment

Retention: This is not a good indicator of the invariant because number of users enroll the free trial dependent on experiments. Good evaluation metrics because it is directly dependent on the experiment.

Net conversion: It is not a good invariant indicator , because it is dependent wih the experiment. But it is a good evaluation indicator, because it is directly dependent on the effect of the experiment, and also shows the change of financial results.

I will look for the gross conversion and net conversion.I think the gross conversion will decrease, it tells us we could reduce our costs. The net conversion might stay in the same

level and shows the change positively influence our revenue. The retention is not necessary for the hypothesis and it requires too many page views as an experiment

## Measuring Standard Deviation

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

Gross conversion: 0.0202

Net conversion: 0.0156

For each of your evaluation metrics, indicate whether you think the analytic estimate would be comparable to 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.

Gross conversion and net conversion both have the number of cookies as their denominator, which is also unit of diversion. The unit of diversion is equal to unit of analysis, so the analytical estimate would be comparable to the empirical variability.

## 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 your experiment appropriately. (These should be the answers from the "Calculating Number of Pageviews" quiz.)

No use of the Bonferroni correction

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.)

60% 29 days

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

This experiment will not influence the already-paid students and has no influence in the Udacity course materials. But for the new students, it will have a effect. Therefor I think a 60% fraction is reasonable.

It is not a very risky experiment 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: [0.4988,0.5012]; observed 0.5006; PASS

Number of clicks : [0.4959,0.5041]; observed 0.5005; PASS

CTP:[0.0812,0.0830]; observed 0.0822 PASS

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

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.)

## Effect Size Tests

For each of your evaluation metrics, compute a confidence interval around the difference.

Did you use the Bonferroni correction? ☐ Yes ☒ No

|   | Lower bound          | Upper bound          | Statistical significance            | Practical significance              |
|---|----------------------|----------------------|-------------------------------------|-------------------------------------|
| • Number of cookies                               | <input type="text"/> | <input type="text"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| • Number of user-ids                              | <input type="text"/> | <input type="text"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| • Number of clicks on "Start free trial"          | <input type="text"/> | <input type="text"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| • Click-through-probability on "Start free trial" | <input type="text"/> | <input type="text"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| • Gross conversion                                | -0.0291              | -0.0120              | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| • Retention                                       | <input type="text"/> | <input type="text"/> | <input type="checkbox"/>            | <input type="checkbox"/>            |
| • Net conversion                                  | -0.0116              | 0.0018               | <input type="checkbox"/>            | <input type="checkbox"/>            |

## 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.)

## Sign Tests

Run a sign test on each of your evaluation metrics using the day-by-data data. Enter each p-value, and indicate whether each result is statistically significant.

Did you use the Bonferroni correction? ☐ Yes ☒ No

|   | p-value              | Statistical significance            |
|---|----------------------|-------------------------------------|
| • Number of cookies                               | <input type="text"/> | <input type="checkbox"/>            |
| • Number of user-ids                              | <input type="text"/> | <input type="checkbox"/>            |
| • Number of clicks on "Start free trial"          | <input type="text"/> | <input type="checkbox"/>            |
| • Click-through-probability on "Start free trial" | <input type="text"/> | <input type="checkbox"/>            |
| • Gross conversion                                | 0.0026               | <input checked="" type="checkbox"/> |
| • Retention                                       | <input type="text"/> | <input type="checkbox"/>            |
| • Net conversion                                  | 0.6776               | <input type="checkbox"/>            |

## 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.

I did not use Bonferroni correction. The Bonferroni correction is a method used to counteract the problem of [multiple comparisons](#) and we are testing every variation. In other words, in this case, ALL criteria need to be satisfied to launch, so it is important to control for false negatives and The Bonferroni correction controls for false positives at the cost of power (less power = higher chance of a false negative)

## Recommendation

Make a recommendation and briefly describe your reasoning.

We need to look at the Gross conversion and Net conversion.

Considering a potential negative impact on revenue, we should not launch the experiment.

For the Net Conversion there has been no statistically significant change, but the confidence interval does include the negative of the practical significance boundary. So that the revenue might be negatively influenced by the change. The Gross conversion indicates that Udacity lowers the costs by discouraging students who have not enough time.

Also I think there shall be further and maybe longer test to decide whether to adopt 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.

I will try to run an experiment with about the Gamification of Udacity course system, including introduce a Badge system and a Ranking system, or even a mini-social network system.

**Null hypothesis** by Gamification of Udacity contents, it will not increase Retention by significant amount.

The Gamification change will be randomly assigned to a Control & Experiment Group. The whole courses for Control Group would not change and remained the same while courses for Experiment Group will have this change

**Unit of diversion** is User-IDs,

**Invariant metric** is Number of clicks, because the user clicks before the experiment, and therefore independent.

**Evaluation metric** is the retention rate, if positive and significant, will show an increase of paying customers resulting from the Gamification.

.

Reference :

[https://en.wikipedia.org/wiki/Bonferroni\\_correction](https://en.wikipedia.org/wiki/Bonferroni_correction)