# 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 metrics: number of cookies, number of clicks

Evaluation metric: gross 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.

Since the changes were made clicking ‘start free trial’ button, any metric before that changes should be invariant.

number of cookies: I used it because it reflects the pageviews of overview page which is before ‘start free trial’ button.

number of user-ids: this is not an invariant metric because the number of users enrolled are counted after ‘start free trial’ button. I didn’t use this as evaluation metric since it is not normalized by the number of click. Percentage of people is important than just the number of enrollment.

number of clicks: I used it because it reflects the unique cookies that click the ‘start free trial’ that become the standard.

click through probability: I didn’t used it but it is still invariant metric because it means number of clicks divided by the number of cookies which are both invariant metrics. I thought above two are enough.

My hypothesis is that this experiment would set clearer expectations for students, 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.

gross conversion: This can be used as evaluation metric and I used it to measure the first step: the percentage of people enroll. I expect gross conversion to be decreased since students with less available time would not enroll the courses.

retention rate: I didn’t used retention rate as evaluation metrics. Although this measures the overall effect of the experiment, it measures by user-id enrolled. This is not only difficult to normalize by the number of clicks, but also hard to measure by day.

net conversion: This can be used as evaluation metric to measure the percentage of people who paid. I want net conversion not to be decreased because I don’t want to lose students who can complete the course. I expect the net conversion not to be decrease.

## 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 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.

Since the unit of analysis of our metrics are the unique number of cookies, and the unit of diversion is also the unique number of cookies to check the pageviews, they are the same.

I assume that the analytic estimate would be comparable to the empirical variability since the unit of analysis and the unit of diversion are the same.

## 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 didn’t use the Bonferroni correction worrying that the experiment would be too conservative. I need 685275 pageviews to check the 1% and 0.75% changes in the gross conversion and net conversion each.

### 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.)

Since I wanted to finish the experiment quite fast, I decided to divert 50% of the total traffic of Udacity. To gain 685275 pageviews that I calculated above, approximately 35 days are needed to finish experiment.

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

I chose to divert all of the traffic. I decide the fraction because I thought this experiment is far from risky. This experiment does not include activity that can hurt someone nor include sensitive personal data. To reduce the resources, I think it is better to run appropriately short. Furthermore, 100% of diversion require approximately 18 days to collect needed pageviews which seemed appropriate length.

# 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, 0.5011) / observes: 0.5006

Number of clicks: (0.4959, 0.5041) / observes: 0.5004

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.**

Since I used gross conversion and net conversion as evaluation metrics, I only needed to check the sanity of number of cookies and number of click. After calculation, both turned out to be sane enough.

## 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: (-0.0291, -0.0120) - statistically, and practically significant.

Net conversion: (-0.0116, 0.0019) - statistically not significant, and practically not significant.

### 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: p value: 0.0026 – statistically significant

Net conversion: p value: 0.6776 – statistically not significant

### 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. Since I want both metrics to match my expectation, I don’t need Bonferroni correction. I expect the gross conversion to decrease and net conversion not to be decrease. Bonferroni correction can be good solution to prevent type 1 error when the any metrics are expected to be significant among many metrics. However, since I expect all of our metrics to satisfy my expectation, I don’t think I’ll need Bonferroni correction.

## Recommendation

Make a recommendation and briefly describe your reasoning.

I expected the gross conversion to be significantly lower and net conversion to be the same. By asking students about their available time, I hope that I can reduce the number of frustrated students while not decreasing the number of paying students. As desired, gross conversion showed statistical and practical decrease. This is good because we reduce the students with low probability of completing the course. The decrease of net conversion is shown as statistically not significant. However, the confidence interval does include the negative of the practical significance boundary. This shows that there is possibility that the decrease can be more than practical significance level. Since there are some risk that the experiment reduce the number of students who pay, I don’t recommend Udacity to implement this strategy.

# 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.

To reduce the early cancelation, I made some hypothesis about why students cancel early. The most probable hypothesis I made was that some students may be disappointed about some of the lesson. Students should have read the courser overview carefully, but some might didn’t or might just didn’t accurately expect some lessons and think they are not appropriate for themselves. As a result, I want to design an experiment that make a small lesson to experience the overview of the course. As a result, my hypothesis is that ‘If I add a class overview lesson before enrollment, the early cancellation of students would decrease’.

The coherent unit of diversion should be user-id since the enrollment, cancellation, and completion occur by user-id. Invariant metric should be the number of student who take overview lesson so that constant and experiment group can be compared with the same number of users. Evaluation metrics to measure the effect can be the number of enrollment by user-id and the enrollment-payment rate(number of pay users/number of enroll users). I expect the number of enrollment by user-id to decrease since more informed user would make better choice. Also I expect the enrollment-payment rate to increase since better informed students are expected to be less likely to make early cancelation.