Experiment Design

Metric Choice

Invariant Metrics:

- Number of cookies
- Number of clicks
- Click-through-probability

Evaluation Metrics:

- Gross Conversion
- Net conversion

- **Number of cookies** was selected was selected as it is a very evenly distributed metric across both the control and experiment groups.
- **Number of user-ids** was not selected as a invariant metric, as we would like to test this change on new visitors, who may or may not necessarily be assigned a user-id yet. This was not selected as evaluation metric, as it did not capture the impact of the changes we were testing.
- **Number of clicks** is another metric that would be excellent as an invariant metric, as it should be the same across both groups.
- **Click through probability** is also unaffected by the change we are making and should be the same across control and experiment.
- **Gross Conversion** is driving to the heart of what we would like to test here. This will help us explore the impact of our change, by evaluating the amount of users who still complete checkout after completing the small quiz about how much time they have to dedicate to the course.
- **Retention** was originally selected as an evaluation metric but unfortunately this data would take far too long to collect for statistical significance as the pool of users is so much smaller than the pool of page views.
- **Net conversion** was selected as our other evaluation metric as it gives us an accurate representation on those who complete the first 14 days and decide that they would like to continue taking the courses.

For gross conversion we will be looking for a decrease, and net conversion for an increase.

Measuring Standard Deviation

Metric	Standard Deviation
Gross conversion	.0202
Net conversion	.0156

Because each of these metrics uses cookies as their denominator, and we are using the cookies as our unit of diversion, an analytical estimate will be appropriate.

Sizing

Number of Samples vs. Power

The Bonferroni correction was not needed, and the amount of pageviews needed will be **685325**.

Duration vs. Exposure

I would divert 50% of the traffic to this change, which would require 35 days to run the experiment. This is the section of the analysis were it was determined that the **retention** metric would simply take too long to track.

The experiment is not necessarily risky, even though it discourages people from signing up if they do not have the time, the amount of people who continue should have a higher conversion rate. This means that the net change should not have much impact, although user satisfaction should be much higher. If it was determined that the experiment needed to be run quicker, a larger amount of traffic could be diverted without much risk.

Experiment Analysis

Sanity Checks

Metric	Lower bound	Upper bound	Observed	Passes
Number of cookies	.4988	.5012	.5006	**Yes**
Number of clicks	.4959	.5041	.5005	**Yes**
Click-through-probability	.0812	.0830	.0822	**Yes**

Each of the sanity checks passes. These will be appropriate invariant metrics just as we thought.

Result Analysis

Effect Size Tests

Metric	Lower Bound	Upper Bound	Statistical Significance	Practical Significance
Gross conversion	-0.0291	-0.120	**Yes**	**Yes**
Net Conversion	-0.0116	0.0018	**No**	**No**

Sign Tests

Metric	p-value	Statistical Significance
Gross conversion	.0026	**Yes**
Net Conversion	.6776	**No**

Summary

I did not use the Bonferroni correction, and the results from the effect size test and sign test do not show discrepancies, although we can now see that net conversion was not impacted as we thought it might.

Recommendation

Gross conversion was negatively impacted by the change just as we expected. Unfortunately net conversion was not impacted and we might want to rethink this design before releasing. By adding the quiz we are discouraging users from signing up if they do not have an appropriate amount of time which will decrease the amount of users who sign up and are unhappy with the product. Without the increase in net conversion though, this change will sacrifice revenue for what is hopefully an uptick in user satisfaction. My recommendation is to explore other designs that outline the expectations and course material but increase net conversion.

Follow-Up Experiment

For a follow up experiment, I would want to test expanding the 14-day trial to a 30-day trial. This would give the students more time to explore the material and get a feel for how long the courses might take them. Nanodegrees have many courses and each course has it's own material and challenges. This would allow them to fully explore all the courses and final projects as well as pace the material.

My hypothesis would be that **Gross Conversion, Net conversion**, and **Retention** would increase with the 30-day trial.

The unit of diversion would be by cookie, although it may be disingenuous to provide a 14-day trial to some users and a 30-day to others. In this case it may be better to offer the 30-day trial as a promotion and compare the data to data from the past.

The invariant metrics would be the same as before, **number of cookies, number of clicks,** and **click through conversion**. We would be especially interested in how many more people were click on the "Start Now" button.

The evaluation metrics would be **Gross Conversion, Net Conversion,** and **Retention**. We would be interested in seeing how many people click on the start now button (which would

hopefully increase with a longer trial time), but also that people would be more inclined to continue past the trial period.

We would be looking for a positive increase in gross conversion and net conversion, but the real confirmation would be an increase in **Retention**.

Sources

https://vwo.com/ab-testing/