UDACITY

PROJECT

Design an A/B test

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Requires Changes

3 SPECIFICATIONS REQUIRE CHANGES

Dear student,

I've reviewed your previous work, what a major progress! There are only a few issues left to be addressed in order to meet requirements, I've left extensive comments that might hopefully prove helpful in addressing those very issues.

Good luck with your next submission and keep up your excellent work!

Metric Choice

A good set of metrics have been selected for the experiment, without missing any necessary or valuable metrics.

Each metric has a clear and well-reasoned explanation of why it was or was not chosen as an invariant metric and as an evaluation metric.

The report clearly states what results we look for in order to launch the experiment and the stated results are aligned with the experiment goals.

Variability

The standard deviations for all evaluation metrics have been correctly calculated.

Each evaluation metric has a clear and correct explanation of whether the analytic variability is likely to match the empirical variability.

Sizing

The number of pageviews given is correct given the students choice of whether to use the Bonferroni correction.

A well-reasoned argument about how risky the experiment will be is made and a fraction of traffic to divert is chosen accordingly.

If you're interested in more details regarding the risk assessment you can refer to lesson 2 sections 4 and 5.
Some hints:

1. Is there a chance that anyone gets hurt because of the duration of our experiment?
2. Are we dealing with sensitive data? (Political attitudes, personal disease history, sexual preferences)
   If not this experiment is not risky and we could divert the entire traffic. (As you correctly do)

The duration of the experiment is correctly calculated given the fraction of traffic to divert that was chosen.

Sanity Checks

The sanity checks have been correctly calculated for all chosen invariant metrics.

The passing or failure of all sanity checks have been evaluated. If sanity checks failed, analysis has been performed to discover why the sanity checks may have failed and the experiment has not been continued.

Effect Size Tests

Correctly calculated confidence intervals have been reported for the difference in all evaluation metrics.

Statistical and practical significance have been correctly reported for all evaluation metrics.

Sign Tests

P-value and statistical significance have been correctly reported for all evaluation metrics.

Results Summary

The report provides good justification for the choice of whether to use the Bonferroni correction.

The answer is correct and so is the rationale in: "Our decision to go or not to go for the modification will be based on having a decrease of the gross conversion AND having the net conversion stable." though it is incomplete. Could you please explain **why** the fact that we need both metrics to meet requirements in order to launch, means that we don't need the Bonferroni correction? Let me go through this once again as this is one of the most challenging concepts involved in this exam. We have multiple metrics and we need all of them to meet some criteria in order to launch. If we had

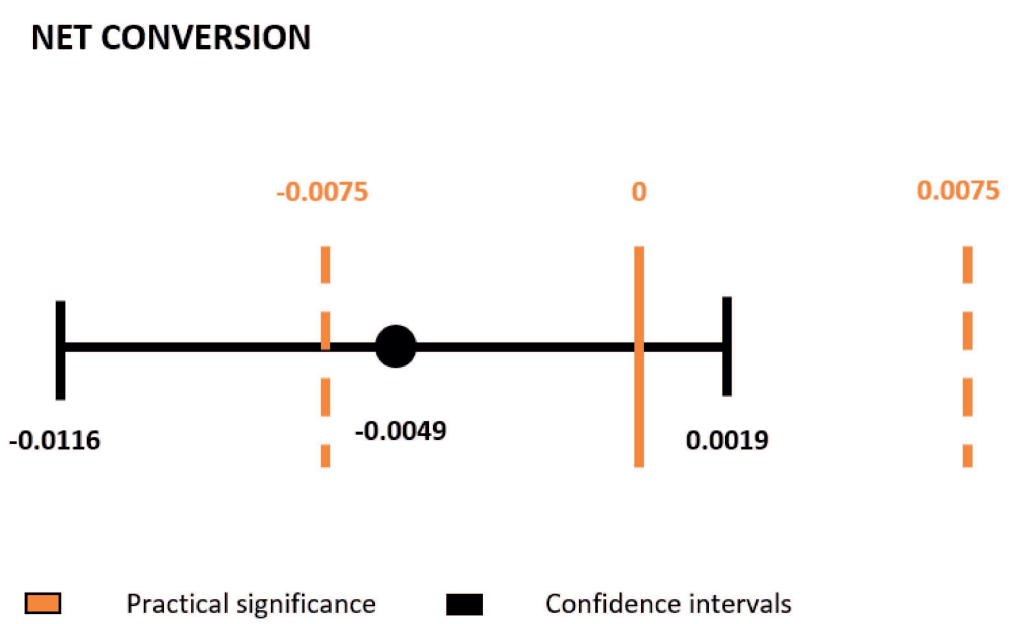
multiple metrics, but we would need only one of them to meet our criteria in order to launch, then the Bonferroni correction would have been useful: In that case we would face the risk that a single metric could meet criteria by pure chance, by mistake. Because we need more metrics to meet some criteria in order to launch, (we need the gross conversion to decrease **and** the net conversion not to decrease), our risks of facing a Type I error is not our main concern. If a single metric demonstrated a false positive, it would not be enough to launch. Our actual risk is that we might reject some metrics by mistake, and that is a type II error. The Bonferroni correction is designed to avoid type I errors when, out of many, it is sufficient that one metric meets the criteria in order to launch. It is useless in cases like ours where, out of many, we need all the metrics to meet criteria in order to launch.

A well-reasoned and plausible explanation for each discrepancy between the effect size tests and the sign tests has been provided.

Recommendation

A recommendation is made that is well-reasoned and supported by the data.

Your answer is correct though part of the rationale is not complete when discussing the net conversion. The problem is not just the fact that we don't have statistical significance for the net conversion, as hinted out in the previous review as well, there is a fundamental further element to be considered : **The confidence interval of the net conversion does include the negative of the practical significance boundary**. That is, it's possible that this number went down by an amount that would matter to the business. Is this an acceptable risk in order to launch? (The answer is straightforward if we consider the second part of our hypothesis)



Follow-Up Experiment

A plausible experiment that would be worth testing has been made. A hypothesis for results of the experiment is clearly stated.

The metrics chosen in the report will be sufficient to evaluate the hypothesis of the experiment, would be possible to measure under most infrastructures, and are well-supported by reasoning in the report.

There is only an issue here: Please note that evaluation metrics are there to inform us about our hypothesis. Please remove any evaluation metric that is not strictly related to informing us about the success of our experiment. Because the hypothesis is concerned with a single goal and has only one part, we need only one evaluation metric.

The report describes a reasonable unit of diversion and gives good support for this choice.

RESUBMIT

DOWNLOAD PROJECT

Best practices for your project resubmission

Ben shares 5 helpful tips to get you through revising and resubmitting your project.

Watch Video (3:01)

Have a question about your review? Email us at review-support@udacity.com and include the link to this review.

RETURN TO PATH

Rate this review

Student FAQ