UDACITY

PROJECT

Design an A/B test

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Requires Changes

1 SPECIFICATION REQUIRES CHANGES

Dear student,

well done improving your submission. There is only an issue left with the Bonferroni correction, please don't be discouraged, the Bonferroni correction is by far the most complex concepts involved in this exam. I'm confident that my further explanations might help addressing the issue.

You're almost there, 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.

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.

There is still seem to be some confusion regarding the Bonferroni correction, in particular regarding the risk we are facing, it is not correct to state that: "In our case, our risk is to accept the modification (meaning to reject the null hypothesis) while the null hypothesis is true. It will have financial impact on Udacity. This is a type I error." The Bonferroni correction is designed to deal with type I errors, if that where our risk we should use it.

I'm quoting my very first comment regarding it as there is no clear way to address the issue: *"To propose your recommendations you will, correctly, consider* both *the net and gross conversion. That is because, in order to launch, you would need them* both *to match our expectations (we look for a decrease in gross conversion and for a no decrease in the net conversion). We are in the situation where more metrics need to be* all *matching what we expect in order to launch. The case where* ***all*** *metrics need to match the expectations in order to launch is not the same as the case where* ***any*** *metric needs to match the expectations. In fact it is the exact opposite:* ***For the former the risk of a*** Type II error ***increases*** *as the number of metrics increases, for the latter* ***the risk of a*** Type I error ***increases.***"

Hope this helps.

Type II error

Type I error

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.

Follow-Up Experiment

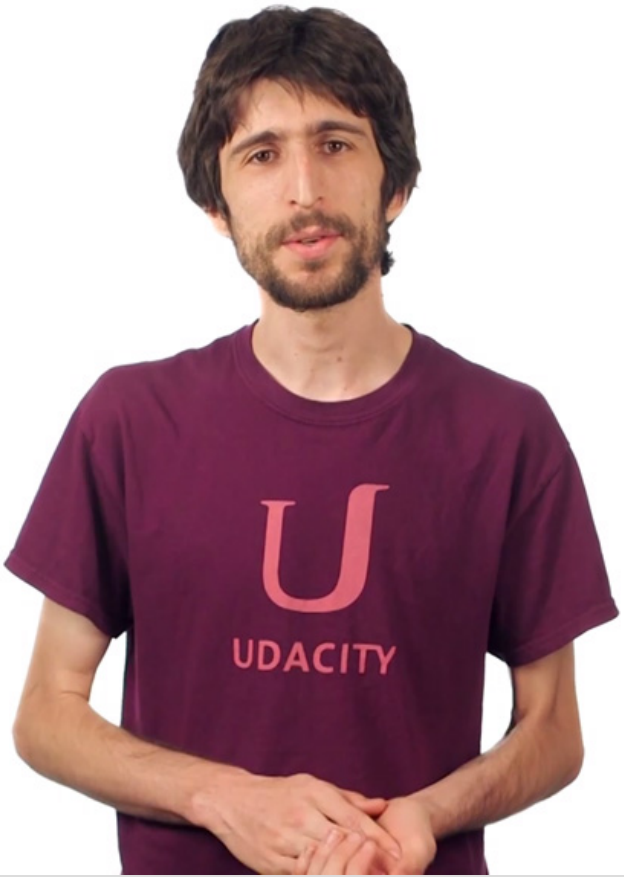
A plausible experiment that would be worth testing has been made. A hypothesis for results of the experiment is clearly stated.

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.

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