UDACITY

Return to "Data Analyst Nanodegree" in the classroom

Analyze A/B Test Results

REVIEW

HISTORY

Meets Specifications

CONGRATULATIONS !!!! You passed this project.

Good links:
https://adespresso.com/guides/facebook-ads-optimization/ab-testing/
https://www.designforfounders.com/ab-testing-examples/
https://www.optimizely.com/optimization-glossary/ab-testing/

Some stats on A/B testing:
https://www.abtasty.com/blog/learn-from-5-ab-test-case-studies/

Khan Academy videos on Hypothesis: https://www.khanacademy.org/math/statistics-probability/significance-tests-one-sample/more-significance-testing-videos/v/hypothesis-testing-and-p-values

OLS Regression: Scikit vs. Statsmodels?

Interpreting Results from Linear Regression

Code Quality

All code cells can be run without error.

Perfect!!

Docstrings, comments, and variable names enable readability of the code.

PART - 1

1. Every thing is fine.
2. To remove duplicate a good way is to use, https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.drop_duplicates.html

PART - 2

When possible, it is always more computationally efficient to use numpy built-in operations over explicit for loops. The short reason is that numpy -based operations attack a computational problem based on vectors by computing large chunks simultaneously.

Additionally, using loops to simulate 10000 can take a considerable amount of time vs using numpy https://softwareengineering.stackexchange.com/questions/254475/how-do-i-move-away-from-the-for-loop-school-of-thought

Fast code:

```
new_converted_simulation = np.random.binomial(n_new, p_new,  10000)/n_new
old_converted_simulation = np.random.binomial(n_old, p_old,  10000)/n_old
p_diffs = new_converted_simulation - old_converted_simulation
```

PART - 3
All Good!!
INTERPRETING LOGISTIC REGRESSION COEFFICIENTS: http://www.juanshishido.com/logisticcoefficients.html

Statistical Analyses

All results from different analyses are correctly interpreted.

The null and the alternative hypothesis are appropriate.

Considering the results of the statistical test (p-value) and the suggested p-critical. Since p-value > p-critical, we can't reject the null. http://www.itl.nist.gov/div898/handbook/prc/section1/prc131.htm

For all numeric values, you should provide the correct results of the analysis.

AWESOME
Getting the stats calculations for both the simulation and z-test correct is difficult at this stage. Great work.

Conclusions should include not only statistical reasoning, but also practical reasoning for the situation.

Spot On!!! Great intuition with the relationship between the different hypotheses statements.

- Part iii is a two-tailed test and Part ii is a one-tail test, can you convert the p-values between each other?
  One-Tailed and Two-Tailed Results
  https://stats.idre.ucla.edu/other/mult-pkg/faq/pvalue-htm/

DOWNLOAD PROJECT

RETURN TO PATH