UDACITY

< Return to "Data Analyst Nanodegree" in the classroom

Analyze A/B Test Results

REVIEW

HISTORY

Meets Specifications

Congratulations Student!! 🎉🎉

YOU HAVE SUCCESSFULLY MET ALL THE REQUIREMENTS OF THE RUBRIC. EXCELLENT JOB ON THIS PROJECT! 😃

- With this project, you have learnt many key-concepts in field of data analysis: A/B testing and its methodology, Hypothesis testing, various statistical measures, z-test, logistic regression classifier and much more.... 👍
- Before you proceed to your next lessons, take pride in the effort you've put into this project. Keep up the great work and effort :)

All the best for your upcoming lessons and projects..!!!
Stay Udacious and Happy Learning.!! 😄

Code Quality

All code cells can be run without error.

Good Job...!! 👍

- All of the code cells run without any errors.

Tips and Suggestions:

Tips and Suggestions:

- You have really done a great job in framing various blocks of code via critical thinking (It is visible from the submission).
- However, since we are using numpy and pandas, we have a faster way to simulate the 10000 trials (Part 2 (h))

- **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. This makes processing very fast**
- Additionally, using loops to simulate 10000 can take a considerable amount of time vs using numpy
  Here is a link with detailed explanation

```
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
```

- Essentially, we are applying the null proportion to the total size of each page using the binomial distribution. Each element, for example, in np.random.binomial(n_new, p_new, 10000) results in an array with values like
  [17262, 17250, 17277...] . This array is 10000 elements large.
- When we divide it by n_new , Python broadcasts n_new for each element and we return a proportion for each element.
- This is essentially is simulating, 10000, the new page conversion rate.
- We do this again for the old page.
- The difference between the two will result in a simulated difference array of length 10000 between the new page and old page conversions.
  Note that this method does not require you to calculate the null values to get the p-value.

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

Awesome Work..!! 😃

- I appreciate that you have organised your code and have taken care of markdown cell and code cells as per relevance. This is a good portrayal of a planned and organised submission!! 👍👍
- Comments, docstrings and appropriate variable names are essential for a good coder.
- These not only guide the viewer through the code but also helps in understanding it easily. You have portrayed these skills well... Keep up this good work in future too... 👍

Statistical Analyses

All results from different analyses are correctly interpreted.

Great job !!🎉🎉

All the results from the different analysis are correctly interpreted.
Getting the stats calculations for both the simulation and z-test correct is difficult at this stage. Great work. !! 😄

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

Awesome job in calculating all the numeric values and provide the correct results of the analysis.....!!🎉🎉

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

Well done...!👍🏽

The results are really good... The submission has a practical conclusion and is easily understandable to a person who has less understanding of statistics.

DOWNLOAD PROJECT

RETURN TO PATH

Rate this review