

XU JOSEPH SHEN

[DATA SCIENCE](#)[DS|CS ARTICLES](#)[MARTIAL ARTS](#)[BLOG 博客](#)[RESUME](#)[LIST](#)

E-COMMERCE AB-TESTING

THE PROBLEM

Company Xero is a worldwide e-commerce site with localized versions of the site. We noticed that Spain-based users have a much higher conversion rate than any other Spanish-speaking country. We hypothesized that one reason could be translation. All Spanish speaking countries had the same translation of the site which was written by a Spaniard. We agreed to try a test where each country would have its one translation written by a local. That is, Argentinian users would see a translation written by an Argentinian, Mexican users by a Mexican and so on.

Here, I made a hypothesis that the local translation treatment will increase the conversion rate.

In this task, I am going to analyze whether the effect of localized-translation is working and interpret the results based on the analysis.

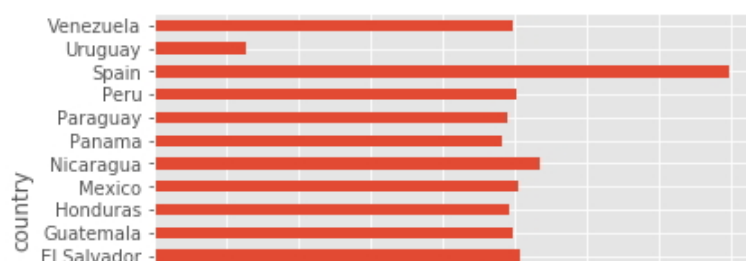
THE SOLUTION

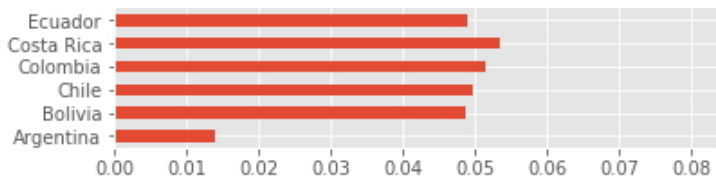
With Python and scipy package, I joined the different tables and conduct independent t-test on different groups. I further explored the confounding variables and explained why the effect is not statistically significant.

PROJECT WEBSITE

THE FINAL REPORT AND DETAILED R CODE CAN BE FOUND [**HERE**](#).

SAMPLE SCREENSHOTS





The Conversion Rate of Different Lational Countries

	statistic	pvalue
Mexico	1.386662	0.165547
Venezuela	-0.562558	0.573740
Bolivia	-0.359969	0.718877
Colombia	-0.799968	0.423733
Uruguay	0.147414	0.882812
El Salvador	-1.154691	0.248251
Nicaragua	0.278873	0.780351
Peru	0.289828	0.771950
Costa Rica	0.401779	0.687863
Chile	1.030324	0.302871
Argentina	-0.990845	0.321766
Ecuador	-0.048257	0.961512
Guatemala	-0.564878	0.572165
Honduras	-0.719713	0.471721
Paraguay	0.146283	0.883702
Panama	0.378121	0.705361

The t-statistics after controlling the country
cofounding variable