Velo.com Revisited Case

.Rmd Template   Data

Make sure to copy these files to the appropriate directory on your computer after downloading!

Case Description

Sara Swan is a data scientist at velo.com, a UK-based online retailer of bikes and biking equipment. She noticed a potential problem with the company's checkout process and recommended the company run an experiment, an A/B test, to determine whether it might be causing lower spending among mobile users. Sarah designed the experiment to run for one month as follows.

- Customers would be randomly assigned to a treatment or control group.
- The treatment group would be directed to the website with the new checkout system, while the control group would get the old website with the legacy checkout system.
- Customers would be presented with the same system for every subsequent visit.

Sarah made some potentially interesting discoveries in her preliminary analysis of the experimental data. However, before reporting her results to velo.com's management team, with her recommendations, Sarah would like to confirm her findings using a different statistical method. She completed her initial analysis using a t-test to compare average spending of customers presented with the old and new checkout systems. Now she would like to do the same comparison of treatment and control groups using simple linear regression.

Do the regression results support her previous findings?

Note to students: The data for this case is identical to the previous case, velo.com. Your analysis for this case should be a *re-analysis* of that data, using linear regression, and should include a *reconsideration* of the results you obtained for the velo.com case.

Appendix. Description of the Data

Rows consist in purchase and spending information during the month of the experiment for velo.com customers who purchased at least 1 item. (Site visits without purchases are not included.) There is one row per customer. Some demographic characteristics of customers are included.

- *customer id*. Unique customer identifier.
- *country*. Customer country of residence.
- *gender*. Customer gender.
- *spent*. Aggregated amount customer spent during the experiment, in dollars.
- *purchases*. Number of items purchased during the experiment.
- *device*. Device most frequently used to make purchases: computer or mobile.
- *checkout_system*. Randomly assigned experimental group: treatment (new) or control (old).

Explore Data

Show [10 ▾] entries   Search: [   ]

customer_id	checkout_system	device	country	ge
1	8968	old	computer	ESP	F

	customer_id	checkout_system	device	country	gender	purchases	spent
2	36687	new	mobile	GBR	F	4	4304.56
3	42232	new	computer	DEU	F	4	396.59
4	82931	old	mobile	MEX	M	4	360.66
5	7010	old	computer	USA	M	4	2597.86
6	83252	new	mobile	GBR	F	4	1458.12
7	51631	new	mobile	FRA	F	4	1703.34
8	38122	old	computer	ESP	M	4	1910.86
9	29745	old	mobile	USA	M	5	1702.61
10	1151	old	computer	FRA	M	2	1937.69

Showing 1 to 10 of 3,483 entries

Previous 1 2 3 4 5 … 349 Next