

A/B Testing Final Assignment

Nathan Cronin

The data utilized for this project was obtained from the public 'dsv1069' dataset on Mode Analytics. This dataset consists of multiple tables detailing users, events, orders, etc. for an online retailer. All actions pass through the events table such as user profile views, item views, and orders, thus it is a very flexible table helping us in our analysis. *All code can be found in the report details section

The purpose of this analysis is to utilize A/B Testing to determine if making an item-level change on website configuration will affect the amount of items that are viewed and

1. Data Quality Check

	item_id	test_a	test_b	test_c	test_d	test_e	test_f
1	2512	1	0	1	1	0	1
2	482	0	1	1	1	0	0
3	2446	0	1	1	0	1	0
4	1312	0	0	0	0	0	1
5	2550	1	1	1	1	1	1

This table only shows the first 1,000 rows.

This all the information from the final_assignment_qa table. Of note, it is not the table we used for our actual analysis, but we did want to determine on what was needed for this table to be analysis ready. It was found that:

This table does not have the necessary information to compute informative metrics.

In it's current state the table is just listing unique key information.

We would need to add values for the dispersion point (date.time) to see 30-day view.

We'd also need binary values showing what users are in control/treatment group.

2. Reformat the Data

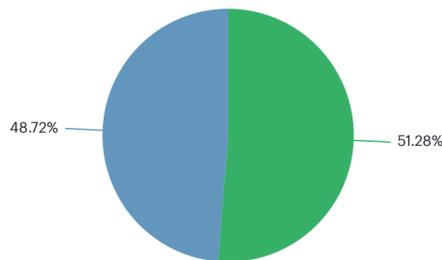
	item_id	test_group	test_assignment	timestamp
1	2512	test_a	0	2015-01-01 mm:dd:ss
2	482	test_b	1	2015-01-01 mm:dd:ss
3	2446	test_b	0	2015-01-01 mm:dd:ss
4	1312	test_f	0	2015-01-01 mm:dd:ss
5	2550	test_a	1	2015-01-01 mm:dd:ss

This table only shows the first 1,000 rows.

3. Compute Order Binary

	test_assignment	items	orders_binary_30d
1	0	1130	301
2	1	1068	286

Binary Order Count per Test Assignment

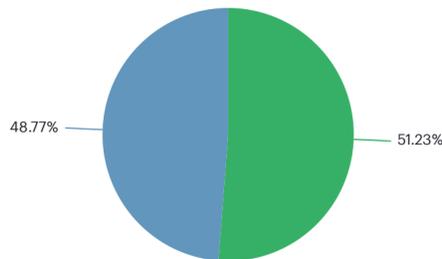


test_assignment 0 1

4. Compute View Item Metrics

	test_assignment	items	views_binary_30d	average_views
1	0	1130	917	1.72
2	1	1068	873	1.73

Binary Item Views per Test Assignment



test_assignment 0 1

Computing P-Value and Lift

Null Hypothesis: Changing the website layout per item does not have an effect on item views and orders

Order Binary Views Binary

P-Value = 0.94 P-Value = 0.72

Lift = 0.53% Lift = 0.73%

As you can see, the p-value is not statistically significant and we cannot say that the change in views and orders was due to our website change. The lift is also negligible, offering no real increase to views and orders. For this we fail to reject the null hypothesis