

Clone of Final Assignment

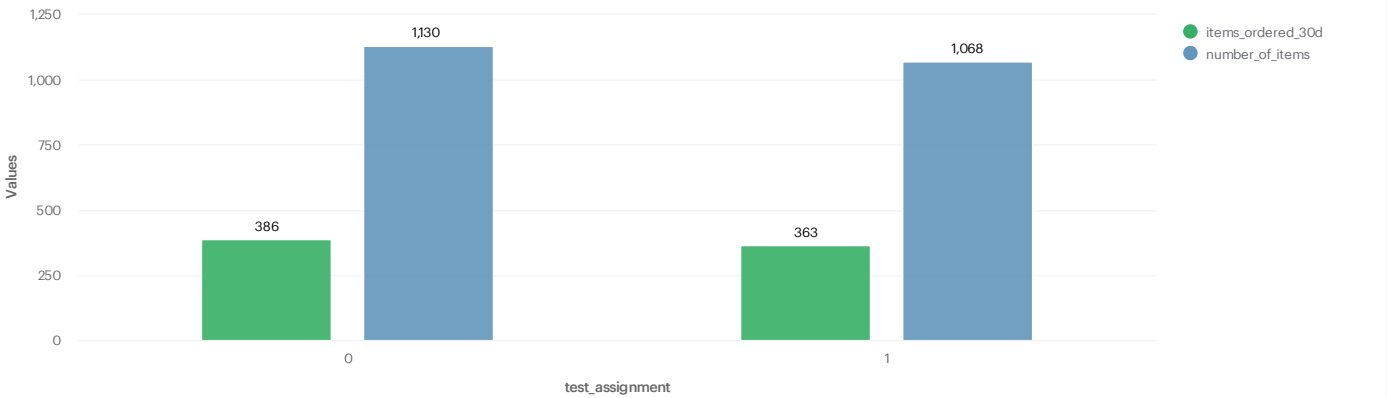
Peer-graded Assignment: AB Testing at the item Level

By Charlotte Fettes

3. Compute Order Binary

	test_assignment	number_of_items	
1	0	1130	
2	1	1068	

Bar chart of number of items per assignment and number of items ordered within 30 days



For order binary, the null hypothesis is that there is no difference in number of items ordered within 30 days between assignment group 0 and assignment group 1.
The alternative hypothesis is that items ordered within 30 days is greater for the treatment group (assignment group 1) compared to the control group (assignment group 0).

Using the <https://thumbtack.github.io/abba/demo/abba.html>, the lifts in metric and the p-value for the 30 day order binary metric using a interval 95% confidence was computed. Results are as follows:

Control:
Success 34%
Treatment:
Success 34%
Lift -0.5%
p-value 0.93
95% confidence interval -12 - 11%

This p-value of 0.93 is very high, therefore, even at a high significance level (5%), the p-value is much higher. Therefore, the null hypothesis is accepted; there is no difference in the number of items ordered within 30 days

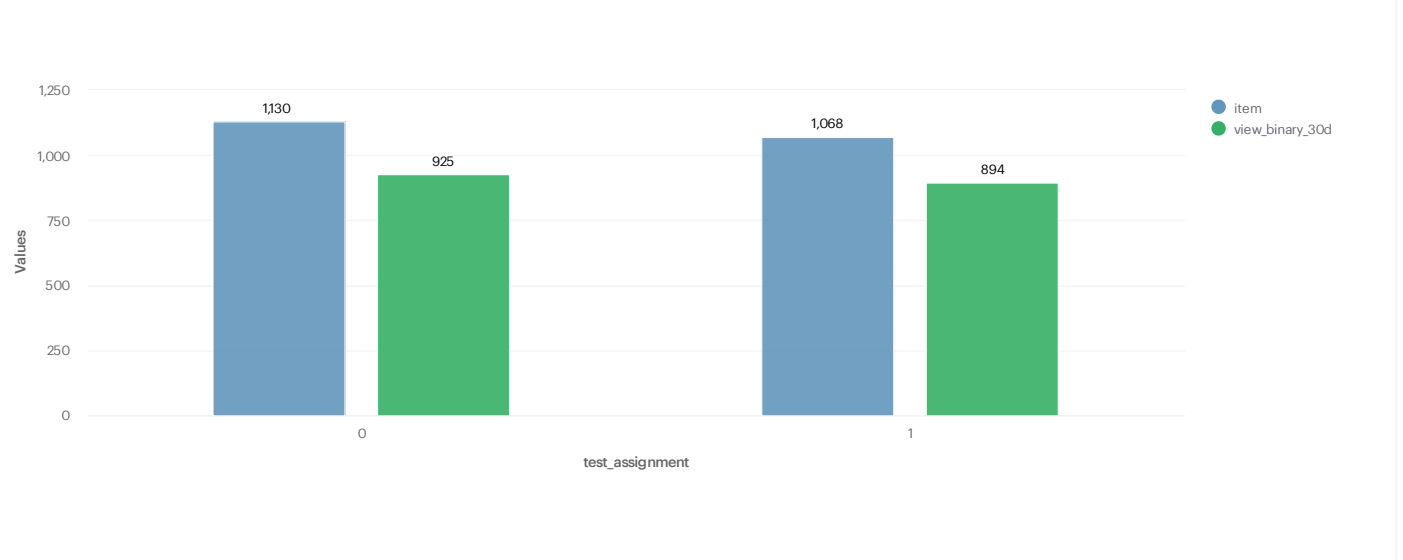
Clone of Final Assignment

between the treatment and control groups. This conclusion is further supported by the 95% confidence interval; we can be 95% confident that the true difference between treatment and control groups in number of items ordered within 30 days is between -12% and 11%. As this interval crosses 0, no difference between groups is a reasonable possibility, therefore we can not be certain that treatment has any effect or if it does, the direction of the effect.

4. Compute View Item Metrics

	test_assignment	test_number	item
1	0	item_test_2	1130
2	1	item_test_2	1068

Bar chart of number of items per assignment and number of items viewed within 30 days



For view binary, the null hypothesis is that there is no difference in number of items viewed within 30 days between assignment group 0 and assignment group 1.

The alternative hypothesis is that items viewed within 30 days is greater for the treatment group (assignment group 1) compared to the control group (assignment group 0).

Using the <https://thumbtack.github.io/abba/demo/abba.html>, the lifts in metric and the p-value for the 30 day view binary metric using a interval 95% confidence was computed. Results are as follows:

Control:

Success 82%

Treatment:

Success 84%

Lift 2.3%

p-value 0.25

95% confidence interval -1.6 - 6.1%

This p-value of 0.25 is still quite high, therefore, even at a high significance level (5% or 0.05), the p-value is much higher. Therefore, the null hypothesis is accepted; there is no difference in the number of items viewed within 30 days between the treatment and control groups. This conclusion is again further supported by the

Clone of Final Assignment

95% confidence interval; we can be 95% confident that the true difference between treatment and control groups in number of items ordered within 30 days is between -1.6% and 6.1%. As this interval crosses 0, no difference between groups is a reasonable possibility, therefore we can not be certain that treatment has any effect or if it does, the direction of the effect.