Final Assignment

1. Data Quality Check

	item_id	test_a	test_b	test_c	test_d	test_e	test_f
	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	3556	1	1	0	1	0	0
6	131	0	0	0	0	1	1
7	1178	1	0	1	0	1	1
8	110	0	1	1	1	1	0
9	47	0	0	1	0	1	1
10	1696	0	0	1	1	1	1
11	3196	0	0	0	1	0	1
12	1578	1	0	0	1	0	1
13	0	1	0	1	1	0	0
14	107	0	1	0	1	1	1
15	487	0	0	1	0	1	1
16	1777	1	1	1	0	1	0
17	1132	0	1	1	1	1	1

This table only shows the first 1,000 rows. View complete results in Report Details.

We are running an experiment at an item-level, which means all users who visit will see the same page, but the layout of different item pages may differ. Compare this table to the assignment events we captured for user_level_testing. Does this table have everything you need to compute metrics like 30-day view-binary?

No, in the table we can't find when the assignment start, we can't find in event table either, so there's an important information missing.

2. Reformat the Data

	item_id	test_assignment	test_number	test_start_date
	item_id	test_assignment	test_number	test_start_date
1	3252	1	item_test_1	2022-01-22 00:00:00
2	642	1	item_test_1	2022-01-22 00:00:00
3	854	1	item_test_1	2022-01-22 00:00:00
4	3095	0	item_test_1	2022-01-22 00:00:00
5	2082	1	item_test_1	2022-01-22 00:00:00
6	3404	1	item_test_1	2022-01-22 00:00:00
7	1805	0	item_test_1	2022-01-22 00:00:00
8	2104	1	item_test_1	2022-01-22 00:00:00
9	999	1	item_test_1	2022-01-22 00:00:00
10	3364	0	item_test_1	2022-01-22 00:00:00
11	487	0	item_test_1	2022-01-22 00:00:00
12	2245	1	item_test_1	2022-01-22 00:00:00
13	215	0	item_test_1	2022-01-22 00:00:00
14	3753	0	item_test_1	2022-01-22 00:00:00
15	586	1	item_test_1	2022-01-22 00:00:00

Reformat the final_assignments_qa to look like the final_assignments table, filling in any missing values with a placeholder of the appropriate data type.

The original format of final_assignments_qa is pivot table, and the format of final_assignments is reverse pivot table. We can convert the original table to a new table with item_id, test_assignment and test_number, but we still lack of time information. We can use the following code to create a table and put some value in event_start_time by default first, and see if we can update in the future.

```
CREATE TABLE IF NOT EXISTS final_assignments
(
 item_id     INT(10) NOT NULL,
 test_assignment  INT(1) NOT NULL,
 test_number    CHAR(20) NOT NULL,
 test_start_date  DATE NOT NULL DEFAULT CURRENT_TIMESTAMP
);
```
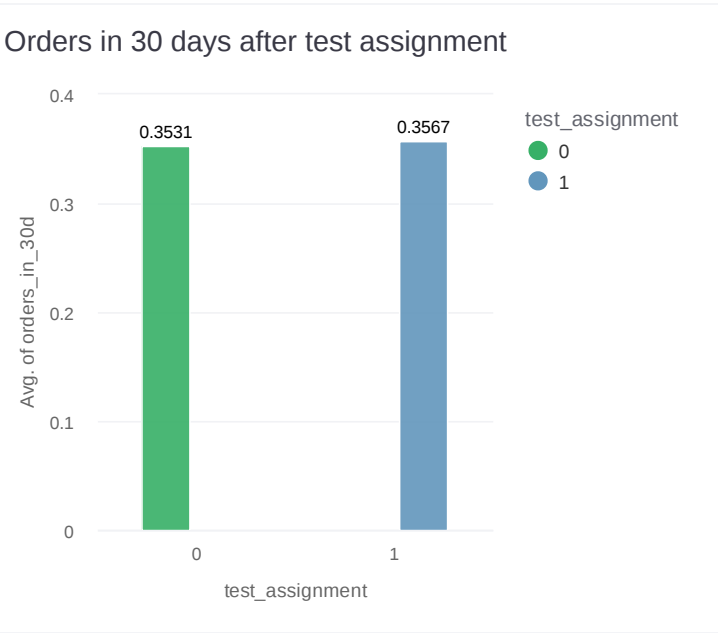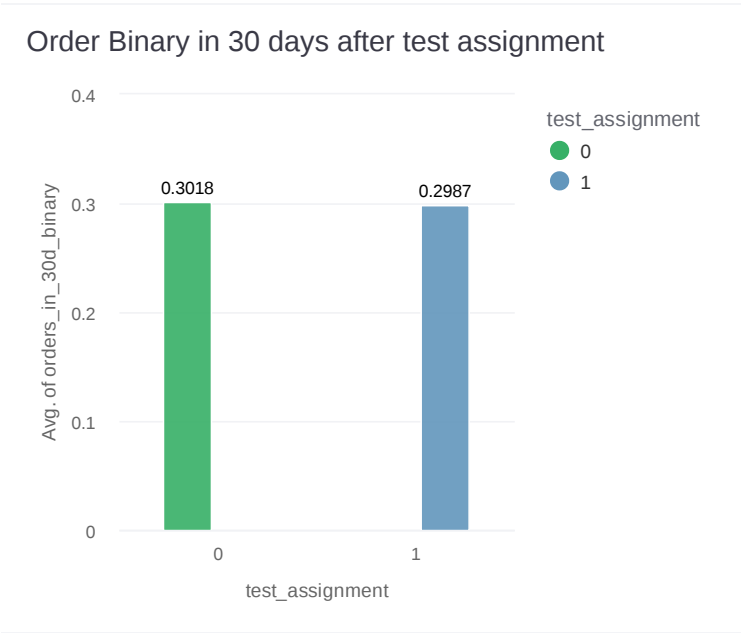
Final Assignment

16	632	1	item_test_1	2022-01-22 00:00:00

This table only shows the first 1,000 rows. View complete results in Report Details.

3. Compute Order Binary

	item_id	test_assignment	test_start_date	orders_i
	item_id	test_assignment	test_start_date	orders_i
1	0	0	2015-03-14 00:00:00	
2	3	0	2015-03-14 00:00:00	
3	5	0	2015-03-14 00:00:00	
4	7	0	2015-03-14 00:00:00	
5	16	0	2015-03-14 00:00:00	
6	17	0	2015-03-14 00:00:00	
7	18	0	2015-03-14 00:00:00	
8	19	0	2015-03-14 00:00:00	
9	21	0	2015-03-14 00:00:00	
10	23	0	2015-03-14 00:00:00	
11	26	0	2015-03-14 00:00:00	
12	27	0	2015-03-14 00:00:00	
13	30	0	2015-03-14 00:00:00	
14	32	0	2015-03-14 00:00:00	
15	36	0	2015-03-14 00:00:00	
16	39	0	2015-03-14 00:00:00	
17	46	0	2015-03-14 00:00:00	

This table only shows the first 1,000 rows. View complete results in Report Details.

From the perspective of average orders, We find that treatment group perform worse than control group, we'll see if there is really statistical significance in section 5.

Also, compared to bar chart with orders in 30 days after assignment, maybe line chart with time in x-axis and include the time before and after assignment could better see if the test really works.

Order Binary in 30 days after test assignment

Orders in 30 days after test assignment

4. Compute View Item Metrics

From the perspective of average views, We find that treatment group perform slightly better than control group,

Final Assignment

|	item_id	test_assignment	test_start_date	views_ir
	item_id	test_assignment	test_start_date	views_ir
1	0	0	2015-03-14 00:00:00	
2	3	0	2015-03-14 00:00:00	
3	5	0	2015-03-14 00:00:00	
4	7	0	2015-03-14 00:00:00	
5	16	0	2015-03-14 00:00:00	
6	17	0	2015-03-14 00:00:00	
7	18	0	2015-03-14 00:00:00	
8	19	0	2015-03-14 00:00:00	
9	21	0	2015-03-14 00:00:00	
10	23	0	2015-03-14 00:00:00	
11	26	0	2015-03-14 00:00:00	
12	27	0	2015-03-14 00:00:00	
13	30	0	2015-03-14 00:00:00	
14	32	0	2015-03-14 00:00:00	
15	36	0	2015-03-14 00:00:00	
16	39	0	2015-03-14 00:00:00	

This table only shows the first 1,000 rows. View complete results in Report Details.

we'll see if there is really statistical significance in section 5. Also, compared to bar chart with orders in 30 days after assignment, maybe line chart with time in x-axis and include the time before and after assignment could better see if the test really works.

View Binary in 30 days after test assignment
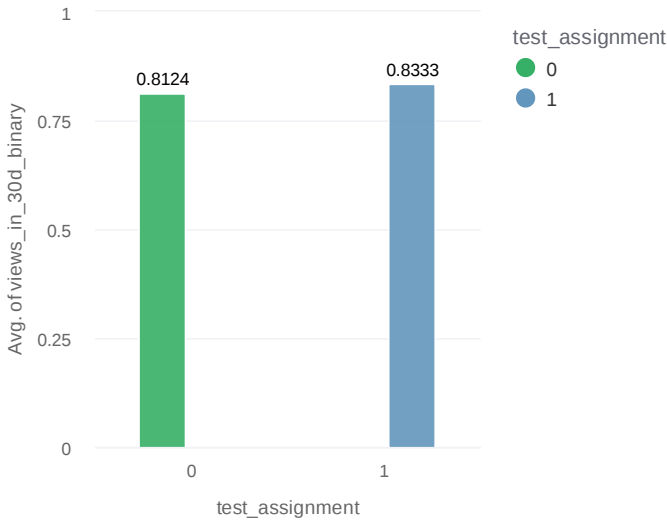
Average Views in 30 days after test assignment
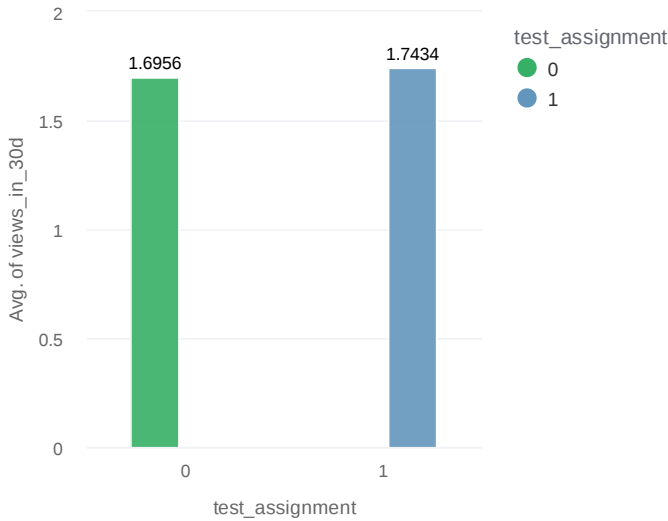
5. Compute lift and p-value

	test_assignment	total_count	ordered_items	orders	vie
	test_assignment	total_count	ordered_items	orders	vie
1	0	1130	341	399	
2	1	1068	319	381	

The information of item_test_2 are shown as left, we can compute lift and p-value to see if the results of this A/B test is statistically significant.

For the order_binary, we got a p-value of 0.88, the confidence interval of lift is [-14%, 12%], there's no different under 95% confidence.

For the view_binary, we got a p-value of 0.20, the confidence interval of lift is [-1.4%, 6.5%], there's no different under 95%

Final Assignment

confidence.

To conclude, we can't reject the Null hypothesis that the treatment group is no different with the control group in both cases. According to the analysis, we can say that considering only whether the item is ordered or viewed after test assignment in 30 days, item_test_2 doesn't work well.