H&M A/B TESTING REPORT

April, 2020

Kathy Zou
Business Analyst

H&M A/B Testing Project

This project will focus on AB Testing, from **data cleaning and traffic chec**k by leveraging **MySQL** to perform the overall **multivariate analysis** and **deep dive** by different categories using **Python**. I used the **Z-Test** for binomial distribution for ratio metrics (**Bounce Rate, ATCR, Checkout Rate, and Conversion Rate**) and **Mann Whitney U tests** for continuous variables (**Revenue**). Based on my analysis, I came up with actionable and insightful recommendations and solutions. Additionally, I provided the next steps based on the current situation.

Introduction

H&M is a multinational fast-fashion company that emphasizes on high fashion at prices significantly below those of their major competitors. Fashion and quality at the best price is the value proposition that they guarantee to their prospective clientele. H&M deals in fashioned apparels, shoes, dress, tops, pants and skirts.

Recently, the product analysis team decided to make some User Experience UX changes on the product page of the H&M website. One of the areas of concern, for example, is that the product analysis team wanted to test whether the product details located at the bottom of the page may be distracting users from clicking the "Add-to-Cart" button. It is for this reason; they came up with ideas of moving the product details to various locations, such as right below the Add-to-Cart button or making the image smaller and moving up the details a bit. After discussions with the product manager and engineers, they designed an A/B Testing (hereinafter called 'Test 1') with two variation groups and a control group, as shown below.

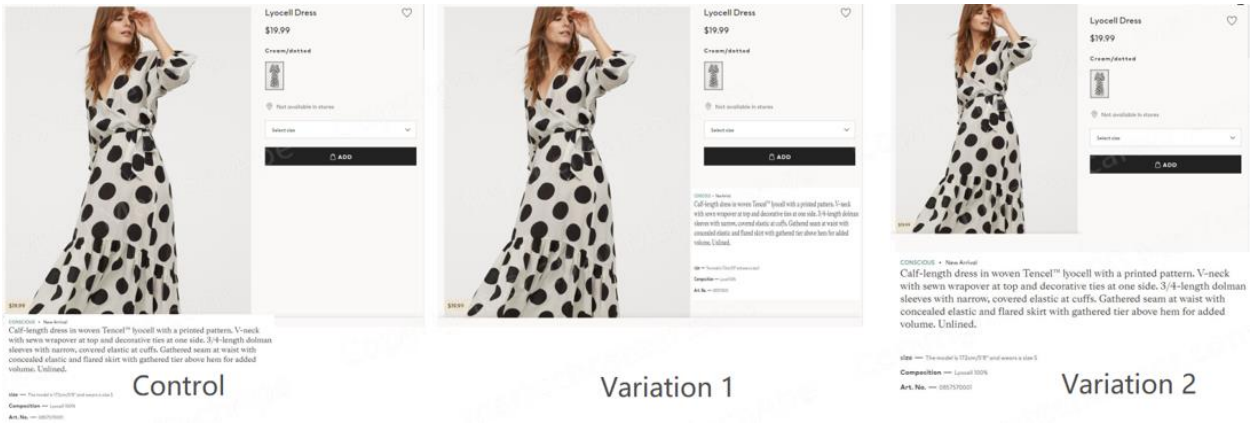

*Figure 1.*Control and variation groups

However, they did not realize that there was another test (hereinafter called 'Test 3') ran on the same page by another group until they ended the test. There was a timeline for the two tests. From the picture below, we can see that both tests started on April 12, 2019. Test 1 ended at May 21, 2019, while Test 3 ended earlier on May 10, 2019.

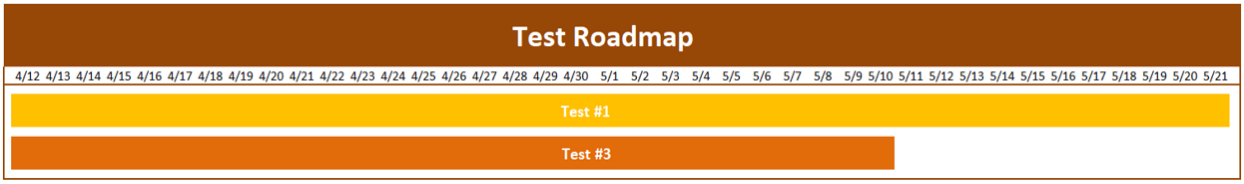

Figure 2. Test Roadmap

After the product analysis team met with the other team, they figured out the content of Test 3, was to add a free shipping banner right above the Add-To-Cart button. They assumed that this change would remind users to realize the shipping fee in advance and consider adding more products on the cart before checkout. Here is the design of Test 3.

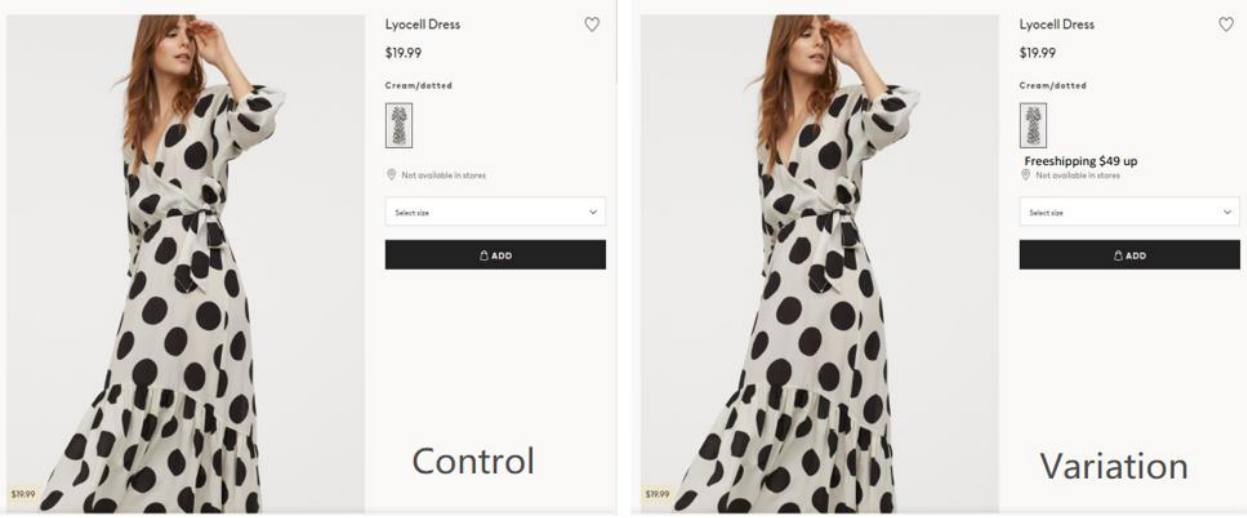

Figure 3. Control and Variation

Executive Summary

- Test 1 performs better when it comes to specific users like customers acquired by email and returning users. **I suggest rolling out variation 1**.
- Returning users react better on test 3. **I suggest rolling out Test 3 on returning users and test the effect when moving the banner to the checkout page.**

Assumption

- The metrics were based on customer and session-level. In my analysis, I checked out both, but I prefer to **use session-level metrics** because it was a fast-fashion company and customers might place multiple orders in a short time when they were influenced by impulse **buying** in making their purchasing decisions.
- The raw data had some issue with sampling, and this implies that some of the test groups **cannot perform deep dive because of the small traffic**. This was particularly the case in the situation at hand as there was an overlap of the two tests, and we should perform **multivariate analysis** by dividing the sample into smaller groups.

- Of all the analysis and calculation on this report, they have all went through the **P-Value trend test**, which means that all the significance results are stable over the time and reliable.

Traffic Flow

Figure 4. Traffic Flow

From the image above, we can see the users will finish the flow before they finally place an order.

* Land on the product page to have more information of the item they choose

* Add the items to cart if they intend to buy them

* Forward to the checkout page and fill all the information to place an order

* Click place-order button and convert

(Users may leave without any actions on the page on the product or cart page)

Metrics

For this analysis, I used the following metrics to track the performance of the test.

- Primary Metrics: ATCR (Add-To-Cart Rate) and Rev (Average revenue for each order)
- Secondary Metrics: BR (Bounce Rate), C/O (Checkout Rate), CVR (Conversion Rate)

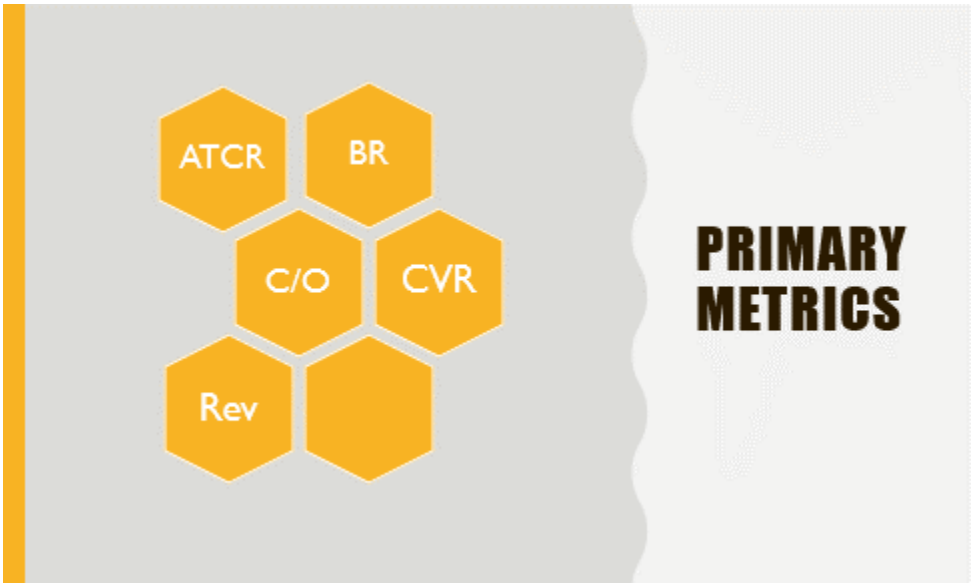

Figure 5. Primary Metrics

Overall Analysis - Session Level

Table 1

Test 1 & Test 3 Results

	test1		test3
	v1	v2	var
BR	No Sig	11.40%	-0.61%
ATCR	0.35%	-11.55%	-1.33%
C/O	No Sig	No Sig	1.38%
CVR	No Sig	-1.19%	No Sig
Rev	-0.33%	No Sig	-2.90%

* Session level

Significant — Positive Lift; Negative Lift

I suggest the rolling out Variation 1 of Test 1. From the table above, we can see Variation 1 of Test 1 has better performance; the ATCR has a significant increase of 0.35%, although there was a decrease in the average revenue. It can be deduced that as the users read the product details, they tended to place their orders more carefully and better informed. There were fewer cases of the influence of impulse spending or buying for each session. I see it as a good sign that our customers have more information about our product. I expect that it would generate higher revenues in the foreseeable future and have brand awareness impacted positively.

Variation 2 exhibits some underperformance than Variation 1. This can be seen in the negative changes in metrics of BR (+11.4%), ATCR (-11.5%) and CVR (-1.19%). Therefore, I think it was not a good and informed idea to move or shift the product details by sacrificing the size of the product image.

As for Test 3, we can see the primary metrics like ATCR and revenue are declining. However, there was an improvement on BR and C/O, which makes sense because the free shipping banner makes users aware of the potential shipping fee for small orders in advance and this is motivated by the customers to add more items on their cart and have a lesser abandon rate for unexpected shipping fee when they checked out.

Test 1 Deep Dive

Cut by Visitor Type

There are four different types of visitors which are new users, email acquired visitors, sign up with no purchase, and users with purchase history. I calculated the performance of each type of visitor. Below are the results that I came up with.

Table 2

Cut by Visitor Type

Test 1	Variation 1			
Visitor Type	BR	ATCR	C/O	CVR
New users	No sig	No sig	No sig	No sig
Email required visitors	No sig	1.03%	No sig	No sig
Sign up with no purchase	No sig	No sig	No sig	No sig
Users with purchase history	No sig	0.43%	-0.79%	No sig

From the table above, I suggested the rolling out of variation 1 except for users with purchase history since there was no negative directional effect on metrics of these groups. For the users who had a purchase history, they had a decline on C/O because they could have been discouraged by the higher value of the cart. Therefore, I suggested that the next step was to launch another test on this group. One of the avenues that I used was, for example, performing user research by sending emails to users who did not check out and find out their reason(s) for the decision that they made.

Cut by Category

I performed the deep dive to see how the users who bought different kinds of categories reacted to the new changes that we had implemented and rolled out. Here is the result:

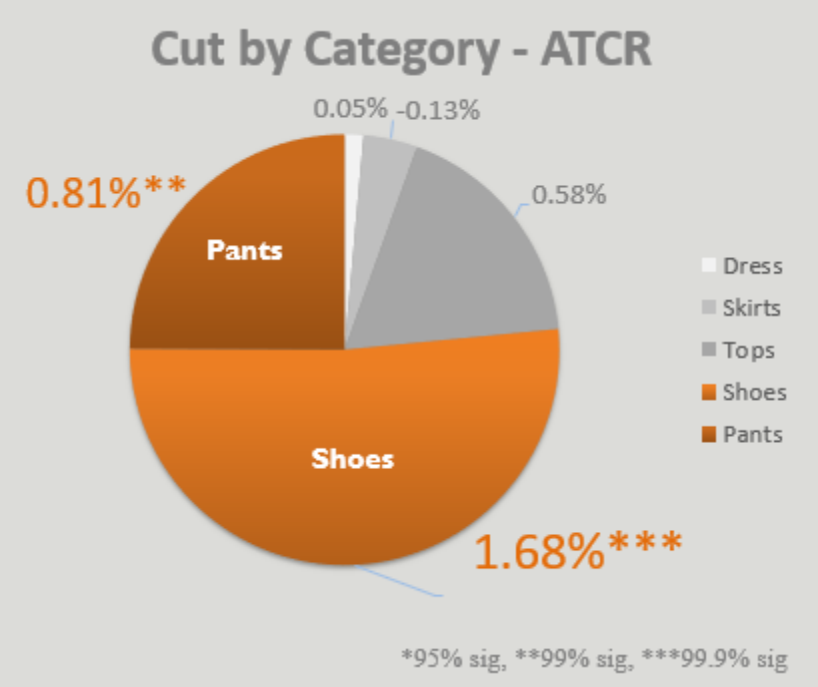

Figure 6. Cut by category

I found out that products like pants and shoes had the ATCR increase since these kinds of categories needed more size and room for detailed information about the products than others for users to make sure they fit. The decision of the users or prospective clients to make a purchase was highly dependent on the specificity and detailed nature of the information available about the product. For the next step, I suggested that we could perform clustering analysis to apply the pattern and see if there were any other categories or users that had the same pattern.

Test 3 Deep Dive

Cut by Visitor Type

Table 3

Cut by Visitor Type

Test 3	Variation			
Visitor Type	BR	ATCR	C/O	CVR
New users	No sig	No sig	No sig	No sig
Email required visitors	No sig	No sig	No sig	No sig
Sign up with no purchase	No sig	-2.10%	No sig	-1.09%
User with purchase history	-1.79%	0.32%	0.01%	No sig

The users that signed up without a purchase went through a decline of ATCR (-2.1%) and CVR (-1.09%). In my opinion, the users may go back to add more products or leave because of the free shipping banner. Moreover, they may tend to abandon the cart for the higher value of the order. When it comes to users with purchase history, I found out that this type of visitors reacted better. Therefore, I suggested the rolling out Test 3 except the signed up without purchase users.

Conclusion

Recommendation

I suggested the rolling out Variation 1 of Test 1 for all of the users. This is because if from a business analytical point of view we are required to be aggressive because it may temporarily decrease the impulse spending, but it may potentially make the users pay more attention to our product details. This is a positive effect that is beneficial to promoting our brand awareness. When we choose to have a lower risk appetite, I suggested, that in this condition, to the rolling out Variation 1 of Test 1 except for the users with purchase history we saw a significant decrease on C/O. Regarding Test 3, I suggested the rolling out it for all the users except the ones that signed up without purchase.

Next Step

For test 1, we could go further and establish the reasons why the returning users were more likely to abandon the cart and check out by using a short and precise questionnaire with open-ended and closed-ended questions. As for Test 3, we could launch a test to check why ATCR and CVR declined for sign up without purchase group. Besides, we could check effects of test 3 when moving the banner to the checkout page or the browsing page to see whether it would have more positive changes on the purchases and by extension the sales volume of H&M.