The 3 key questions to ask after every in-store test

Kaitlyn Choi
PhD, Data Scientist, MarketDial

Introduction

When we think of A/B testing, we often think of the hypotheses, the methodologies, and the results. But none of these things have much meaning unless they are fully synthesized into actionable data after the test is completed.

Knowing what the results are is different from knowing what the results mean. For example, test results might:

1 Highlight positive lift in one geographic region but not another.

2 Predict that rolling out the initiative across the entire fleet is less cost effective than rolling it out to select stores.

3 Show an increase in sales on one product but a decrease in sales on another basket item.

What does this mean to you as a company? What is the best solution or course of action based on the knowledge you now have?

The end of your test produces an accurate result, but that is not the end of the process – the end of your test period is an exciting beginning to further analysis. By thoroughly comparing and analyzing the impact of the data, you will be fully equipped to make well-informed, valuable decisions.

# Finding the results within the results

Your A/B test is complete. You know the surface-level results. Now what?

While test results can readily show you whether an initiative yielded the lift you expected, they also generate multiple follow-up questions.

To fully understand where your opportunities are greatest, three essential questions to ask following every test are:

1 **Why** did the initiative impact the lift — what is driving the impact?

2 **What if** I roll out the initiative to the whole fleet — what would be the overall impact?

3 **What else** is influenced by the initiative — how does it impact correlated items?

All it takes is a click on an analysis tool to get deep-dive answers to these key questions. That means within minutes you can expand your insights to know what's driving the lift, what the overall rollout impact will be, and how the initiative impacts correlated items.

Why?

What if?

What else?
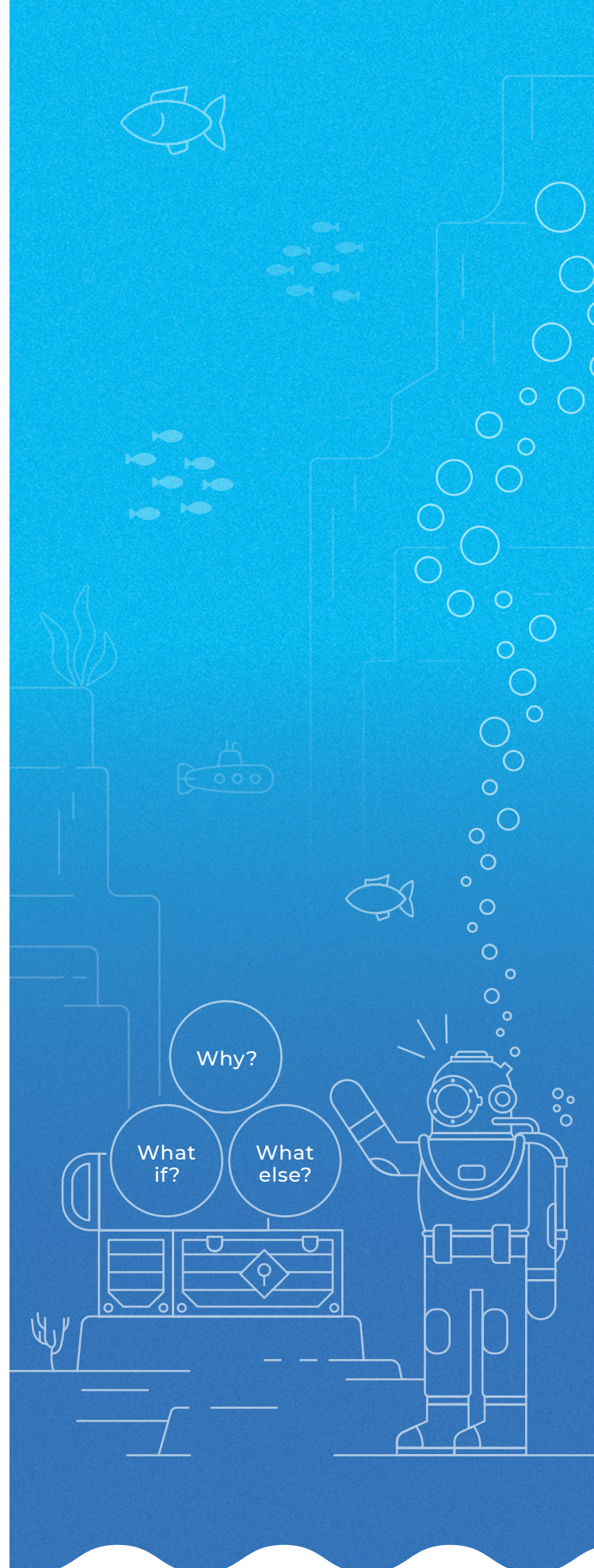
Driving the impact:

Why did the initiative impact the lift?

Test results can show where the lift was highest and where it was lowest, but once that has been determined, it's time to extrapolate why it was higher at some sites and lower at others. What is driving variations in the lift across test sites?

A/B testing features multiple sites that — while part of the same fleet — are different from one another in their own attributes, such as:

- location,
- store size,
- store type, and
- socioeconomics and demographics.

These variables mean that a test result is likely going to be different across the tested sites. Even if a test shows a positive lift overall, some sites may show a negative lift, and it is important to pinpoint which attributes are driving the disparity.

Asking why one site was positive and one was negative reveals the true reason for the lift. When you know what specific attributes make up the positive result (e.g., your initiative worked well in a wealthy neighborhood), you can better understand your initiative — which customers it appeals to, which store size or age it works better with, etc. These insights can help you refine your initiative and/or generate new ideas.

A good driver analysis will:

1 Divide test sites into two groups — one with the positive % lift and the other with the negative % lift.

2 Determine which site attributes are associated with each group and determine how strong the association is.

Case in point:

A convenience store retailer ran a test where they implemented new drink machines at their stores and evaluated how they impacted the sales of hot food products. The test result was overall positive – the revenue of the hot food products increased. A driver analysis on this test revealed that specific site attributes such as median age and income per capita were correlated with the positive lift. Consequently, the client could better determine if they wanted to put the new machines in every location or if it was more financially prudent to only place the machines in sites with that median age and income.
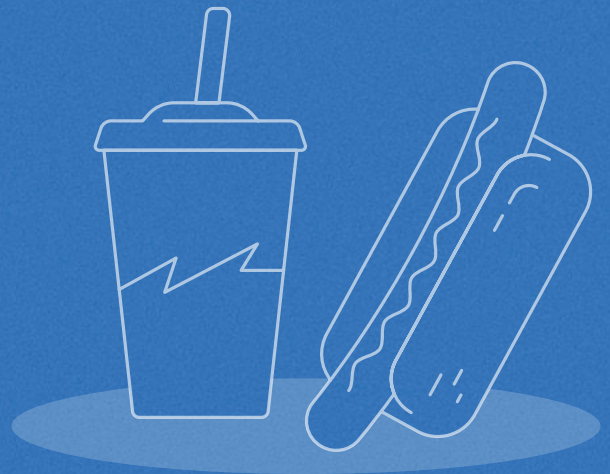
Evaluating the overall impact:
What if we rollout the initiative to the entire fleet?

While site attributes are key to both the driver and impact analyses, the driver analysis focuses on which attributes drive the outcome, whereas the impact analysis focuses on the financial impact those attribute variations will have after rollout.

Once a test is completed, the next step is determining how or if to roll out the initiative. To calculate the potential return on investment (ROI), it is important to estimate the revenue from rollout. That's where impact analysis comes in. A strong impact analysis will:

◇ Define each store in the test according to its attributes, evaluating how different attributes are related to the lift.

◇ Use those learnings to predict the likely lift for each of the stores not tested.

◇ Compute both the full rollout revenue and the optimal rollout revenue which includes only the stores with the positive estimated impact.

For example, an impact analyzer may determine, with 85% confidence, that rollout across the fleet will result in $450 million in annual revenue, whereas rollout across only the stores designated with positive-lift attributes will result in $500 million – eliminating the $50 million in losses that would occur if sites with negative-lift attributes were included in the rollout.

An impact analysis estimates the financial impacts of rolling out the initiative only in stores with positive lift versus rolling it out across the entire fleet.
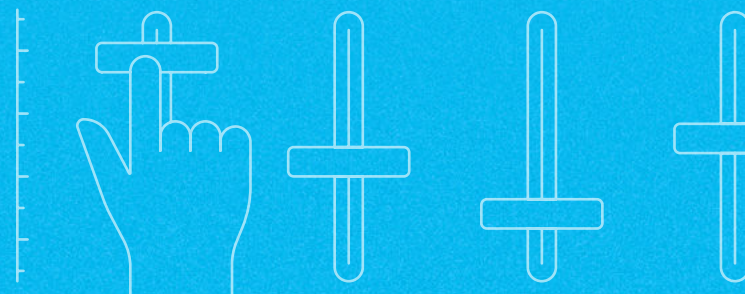
Correlating the impact:
What else is influenced by the initiative?

The lift percentage only measures how much sales increased or decreased on the products tested. But this leaves out an important insight: other items may also have been impacted. It is important to analyze all the transactions that include the tested or "target" product to see if other items in the "basket" have been affected.

Suppose you run a test built to determine the effect of a lower price for denim jeans and learned that denim sales increased with a 5% lift in sales. Great! But wait. You run a basket analysis and discover that graphic T-shirt sales decreased with a 7% decrease.

Knowing the lift percentage is valuable, but it's far from the whole story. By understanding what's driving the lift, what the financial impacts are for different rollouts, and how other basket items are influenced by the initiative, you will be well-equipped with all the information you need to make the best business decisions.

Or, you run a promotion of beer and a basket analysis tells you that pizza was bought together with beer more frequently during the promotion than before the promotion.

A basket analyzer gives you a look inside customer baskets to learn their contents and how the composition changes over time.

You choose the target products, the date range, and the stores to analyze. A basket analyzer finds the transactions that included the target products and exposes:

◇ How much of the basket was made up of the target product(s) vs. other products.

◇ The top product categories purchased together with the target product(s).

◇ How the basket composition changed over time.

With this tool, you can understand what other products were also influenced by your promotion or price change. You can also focus on a different product or alter the date range to see how purchasing behavior changed and what that could mean for your initiative.

Getting deeper insights to maximize business wins

For any A/B test you conduct, the driver, impact, and basket analyzers make it easy to supercharge your findings. With MarketDial, these tools are intuitive, fast — even fun. Your teams are empowered to test rapidly, regularly, and precisely. This is how you achieve more, sooner than you thought possible, powered by our platform and partnership.

Why MarketDial

More than 100 leading companies and global brands rely on MarketDial to help create a culture of experimentation with intentionally easy A/B testing. MarketDial's simple interface makes asking, "What if?" the cost-effective protocol for intelligent action. **Learn how now**

About Youngeun (Kaitlyn) Choi, PhD, Data Scientist, MarketDial

Kaitlyn's biology degrees are from Bard College (B.A.) and Harvard (PhD). Her passion for learning led her to become a professor at Georgetown University, publishing a book on developmental biology and co-writing and starring in a Korean educational program, also on developmental biology. When these challenges no longer gave her that thrill of discovery, she focused on an even better fit — revealing the power of data science.

Her role at MarketDial gives her the ideal forum to delve deep into data and then translate her insights into client business wins. That's not the only thing she translates; as a South Korean native, she thinks in English, dreams in English but still counts in Korean. When she's not inspiring people with data, she's playing tennis, snowboarding, and learning the viola.