MarketDial

White Paper Series
Control Selection

The power of precision

Mastering control selections for in-store testing

Contents

Introduction

Is change measurable? The answer is yes. In science, change is measured by comparing a treatment group that implements a change against a comparable control group that remains unchanged. The key word here is *comparable*.

In A/B testing, the closer a control group matches a treatment group, the easier it is to reduce noise and identify how a specific change impacts outcomes. Therefore, for control selection, the goal is to select another site that is as similar as possible to the treatment site in both sales and demographics.

Without a control site, analysts can't be sure if the results they see stem from the implemented change or from other irregularities. Choosing comparable control sites is therefore essential to creating a baseline to help determine if changes have a distinctive impact.

Comparability

In science, a hypothesis defines what is trying to be proven. The same is true in business. A strong hypothesis identifies what, specifically, is being examined as a potential solution to a business problem. It builds the foundation of an effective test.

Control sites are matched to treatment sites based on two primary factors:

- The historical correlation over the primary metric of interest, such as revenue.
- Their demographic information, such as average household income

These sites can be selected by weighting the importance of both these factors. For example, 80% of the selection can focus on historical correlation over the primary metric and 20% can focus on similarities between demographic variables. Each treatment site is then matched to a possible control site and given a comparability score based on these factors and percentages.

Custom variables can also be introduced depending on the test. If the test has a variable specific to just this initiative, that variable should also be considered in the control matching.

When choosing a control site, the goal is to compare apples to apples, not apples to oranges.

Balancing variables for accurate comparisons

It's much easier to measure how two c-stores with similar trends and demographics perform against each other than it is to measure how a low-performing c-store in rural America performs against a high-performing c-store in downtown New York City.

Note that control matching is not always intuitive. Two sites that may not seem comparable to the naked eye may actually have similar sales patterns. In most cases, the sites don't necessarily need to be in the same state or have the same layout in order to be comparable in their sales patterns.

Generally, the higher the comparability score, the more accurate the results will be – the more accurately we can capture the impact of the changes made to the treatment sites. Likewise, a lower comparability score makes it more difficult to identify cause and effect and to measure the change impact. Therefore, the control site that has the greatest comparability will be matched to the treatment site for the test.

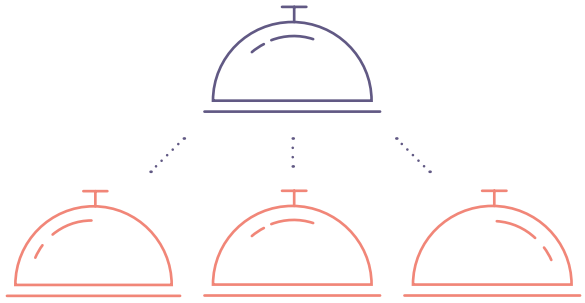
In a similar vein, a group of treatment customers can be matched to a group of control customers. These groups should have similar distributions of key characteristics. For example, if 5% of the customers in the treatment group are in California, then 5% of the customers in the control group should also be in California.

One-many control matching

While it is most common to match one treatment site to one control site, occasionally a higher comparability score can be obtained by matching one treatment site to multiple control sites.

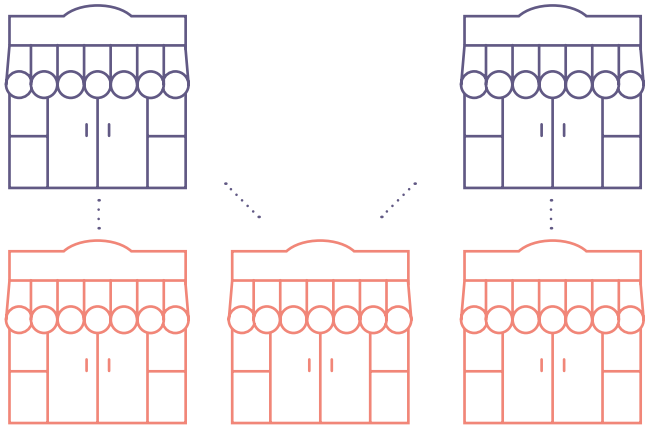
Sometimes unexpected challenges arise at a control site mid-test, such as store closures or staffing challenges. These unforeseen circumstances can inadvertently skew test results. If a control site closes, for example, and starts reporting no sales, the data will suggest the treatment site is outperforming the control site, reporting significant lift.

One-many control matching can protect against these irregularities, averaging them out with other, unaffected sites. Selection is the same as one-to-one selection where each site is given a comparability score and matched accordingly, the primary difference is that multiple sites are chosen for each treatment site as opposed to just one control site per treatment site.



A restaurant chain facing ongoing staffing challenges opts to have three control sites for every treatment site. After evaluating comparability scores, they select the three sites with the highest reported score for each treatment site.

Repeat Controls



A grocery store wants to run a test in 50 stores but only has 100 stores in its fleet. Because there are only 50 stores to choose from for control sites, and because they want high comparability for the 50 treatment stores, they opt to use repeat controls.

Similar to how it can be helpful to match one treatment site to multiple control sites, it can also be advantageous to match multiple treatment sites to one control site.

Many treatments to one control can be useful for retailers with a small fleet size who may have difficulty finding enough comparable control sites to match with each treatment site. The challenge with this format, however, is that if anything disrupts that one control site, it will skew results at multiple treatment sites.

Selection for repeat controls is the same as other selections – control sites are chosen after calculating the comparability score for each treatment-control pair until the highest comparability score is found. However, the difference is that once a control site gets matched to a treatment site, it will then cycle back into the control pool and can be chosen again as a control site based on its comparability score in relation to other treatment sites.

Automating control selection to enhance precision

At MarketDial, we understand how vital it is to select accurate control sites. That's why we've created an app that can automate the process for you.

MarketDial employs advanced algorithms and statistical analysis to automatically select control groups that are representative of the treatment groups. This automation reduces manual intervention and potential biases in the control group selection process.

Why MarketDial?

MarketDial provides retailers with the tools they need to validate initiatives quickly and confidently. Our mission is to maximize our clients' potential by offering accurate testing solutions and actionable insights. With our automated analytics and in-store testing tools, businesses can trust they are making data-driven decisions that drive success.

Want to learn more? Contact us at explore@marketdial.com