Case study: In-store display and visual merchandising analytics

Received (in revised form): 3rd April, 2020

Gary Angel
Chief Executive Officer, Digital Mortar, USA

Gary Angel is the Founder and Chief Executive Officer of Digital Mortar. Digital Mortar provides comprehensive collection and measurement of the retail customer journey. Following Ernst & Young's acquisition of his previous venture, Semphonic (the country's leading digital analytics consultancy), Gary led EY's Digital Analytics Practice. He is the author of 'Measuring the Digital World' (FT Press, 2016).

Digital Mortar, 1005 A Street, #211, San Rafael, CA 94901, USA
E-mail: gary.angel@digitalmortar.com

Abstract In this project, video and advanced machine learning were used to analyse shopper behaviour in a key display area at multiple locations of a multi-billion-dollar retailer. The goal was to understand the volume of shoppers using the area, engagement with the displays, and whether the displays generated product interactions and takeaways. In addition, the system was used to design and support an aggressive testing programme to optimise the display area. Measurement answered the usage questions and revealed obvious opportunities for improvement. Structured testing revealed that the geometry of the area heavily impacted usage and engagement, that shopper flow was strongly influenced by changing the density and alignment of the features, and that there were opportunities for improving product mix and layout. As the purpose of a store is to get shoppers' eyes and hands directly on product, the ability of product displays to attract and engage shoppers is critical to retail success. Like many aspects of physical retail, however, merchants have little visibility into the success of any given display and insufficient measurement to drive testing and improvement programmes. This case study shows how measurement and testing in display has become possible using people-counting technologies.

KEYWORDS: retail analytics, store analytics, merchandising, display, endcap, display measurement, display performance, display engagement

BACKGROUND

Display matters. How and where a product is displayed in a store make a substantial difference in its sales volume. Key areas of the store (near cash-wrap, endcaps, etc) will have many more shoppers pass through than will less central or heavily shopped locations. Taking advantage of (and steering) that traffic is the goal of visual merchandising.

To that end, most retailers use and experiment with a wide variety of display types. With almost any display, there are inevitable trade-offs in terms of the number of products displayed, the degree to which each product is exposed, the visual appeal of the overall display, the footprint of the display and, of course, its cost.

With an almost infinite number of possible strategies for any given display, a store's display merchandising strategy is one of the key factors driving store success.

In this particular case, the retailer was focused on a feature area usually reserved for displaying trending styles. In most store layouts, the feature area was near the front of the store entrance and consisted of multiple display tables.

With little existing information, the initial goal of the project was to collect enough data to baseline performance and answer a set of basic questions about usage and success. Once a baseline was established, it was hoped that the performance and usage data would suggest potential tests that might be run to improve performance and that the system could then be used to determine the success of each iteration of the programme.

Three stores were chosen for the project. The project began in September 2019 and, at the time of writing, is still running.

EXECUTION

Part 1: Shopper measurement technologies

It is impossible to get a deep understanding of display performance just by using downstream metrics from sales. Those metrics will, at best, provide indications of overall success. But they will not give many clues as to how improve a poorly functioning display and they are often driven by variables that have little or nothing to do with actual visual merchandising performance.

To really understand display performance, it is essential to be able to measure real-world opportunity and interactions. In other words, one must be able to measure actual shopping behaviour.

For this project, a camera and advanced single-board computer (see Figure 1) were installed at each location. Initially, two matrixed cameras were considered, but a single camera proved sufficient to provide coverage for the feature area. The cameras were mounted in the ceiling (roughly over the centre of the target area). Installation was straightforward and was done in the morning before the stores opened, thus removing the necessity for overnight or additional onsite security. The camera directs a video feed to the computer, which uses advanced machine learning to process the video stream, identify people, track them through the space and identify display interactions. The resulting data stream is nothing more than a random, unique person identifier (there is absolutely no personally identifiable information (PII) or facial recognition involved in these system), a timestamp, x,y coordinates and a confidence level about product interactions and take-aways. No video leaves the store or is retained anywhere. With some deployments, height and gender can also be captured but neither was used for this implementation. The system is as compliant with PII regulations as, for example, a door counting system.

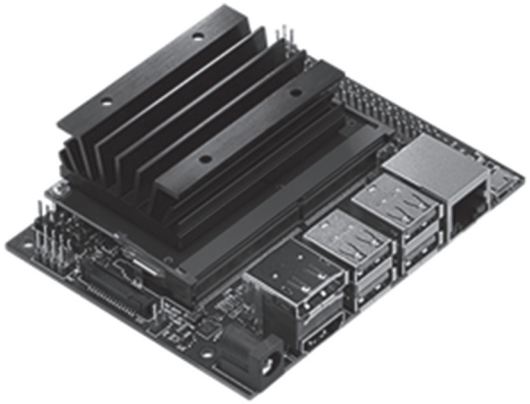

Figure 1: Sample single board computer

While video is captured at 30 frames per second, data are aggregated up to a time resolution of around one second. This means the shopper's true position is known on a second-by-second basis.

Putting these two devices together provides rich, detailed metrics about almost every aspect of shopper behaviour at or near a display. The components are off-the-shelf, highly reliable and not particularly expensive. Together, they provide a complete technology solution for collecting the data necessary to measure and optimise displays.

Part 2: Translating date to metrics

In combination, the camera and single-board computer capture very basic data. The data contain no PII, which is great, but taken in isolation, this kind of data is not directly usable. Like most internet of things (IoT) data, what is obtained is a massive stream of detailed behavioural events. Each event looks like this:

> Random Shopper ID, TimeStamp, Height, X Coord, Y Coord, Interaction Coord, Interaction Confidence

A shopper passing through or by a display might generate 5–10 of these events. A shopper that engages with a display may well generate hundreds of these events. With hundreds or thousands of shoppers passing by a display in a day, the system will generate tens or hundreds of thousands of individual data points.

To make such data useful, the data must be sessionised, mapped onto the store and translated into useful key performance indicators (KPIs).

The first step is to sessionise the data by Shopper ID. Unlike full store shopper journey tracking, the shopper ID is only good for as long as a person remains in view of the camera. This means that if a feature display is located in a common pass-through area, the same shopper may be counted several times as a unique individual. From a 'visits' perspective, this is fair game. If a shopper passes a display twice, there are two opportunities for engagement. However, it does make comparisons with door-count less than useful.

Step two — and this is truly the critical step — is to map the x,y coordinates to the area. This is typically done down to a square-foot level. The area must be mapped with each display placed accurately on a digital planogram. The digital planogram builder (see Figure 2) used allows the analyst to construct multiple digital layouts (using a point and click interface on top of a daily snapshot of the target area), which can be seamlessly swapped into any analytics view. This allows for reporting at the area, display, face and edge level.

One of the real challenges to building ongoing display measurement is the potential (and pace) of change in the store. Modern displays are often quite mobile — and with a table on wheels, it is quite easy for a

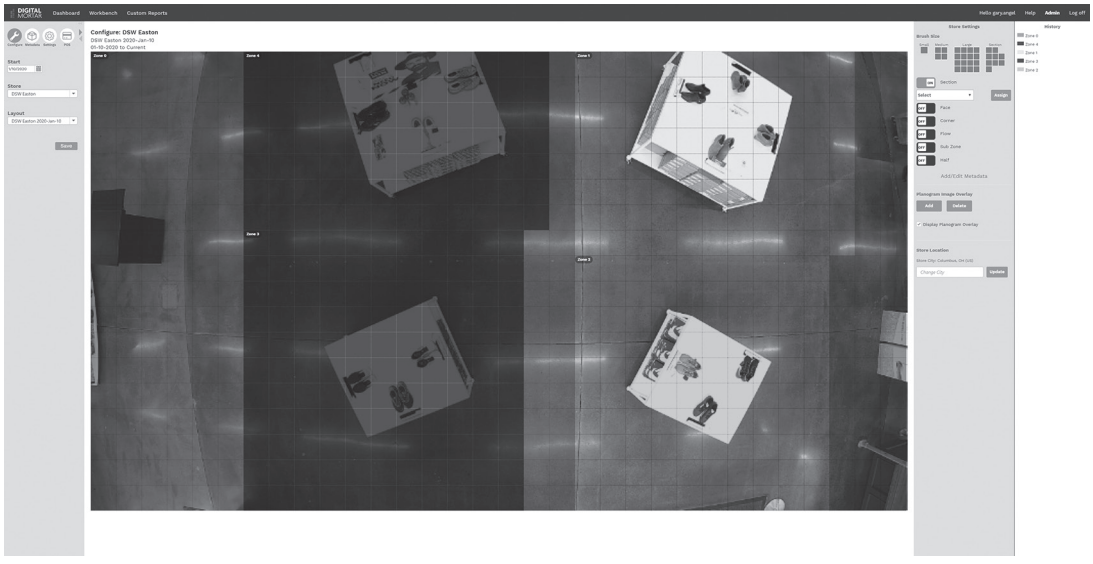

Figure 2: Sample digital planogram layout

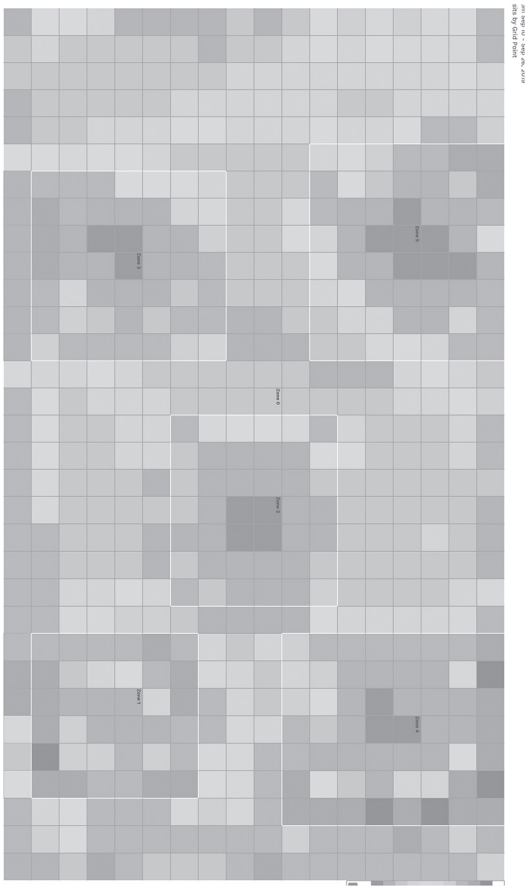

Figure 3: Sample traffic flow — feature area

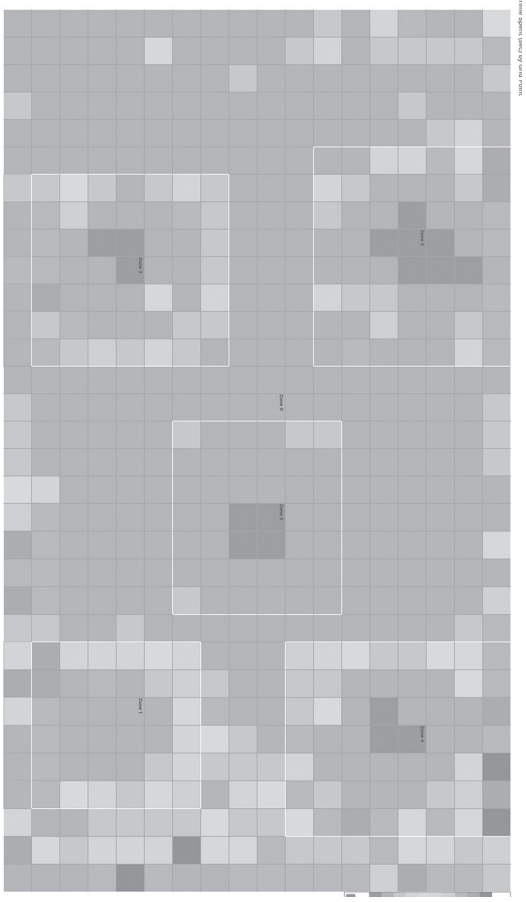

Figure 4: Sample shopper time — feature area

display to shift position unintentionally. In the project, this proved to be a significant challenge. During the early phases of the rollout, displays were getting moved on an almost daily basis.

To counter this, the system starts each day by taking a snapshot of the area before the store opens and then mapping the newest image against the prior one to detect any changes in layout. Of course, this also necessitated some back-and-forth with store managers to make sure that the area remained stable at the macro level during tests.

Part 3: Developing the baseline

The first analytics goal of the project was to measure each location to answer the initial questions about usage and performance.

There was considerable interest in how many people passed through the area and — even more — how many shoppers actually engaged with the displays. Maintaining feature displays is work; besides, it is not worth running tests on an area with insufficient traffic to support analysis.

The initial tests identified that the selected areas in two of the locations garnered plenty of traffic. Visit counts ran well above door counts (meaning that a typical shopper passed through the area between one and two times during a visit). At the third location, the area selected for measurement was in a corner of the store and the traffic volumes were much lower. Because of this, testing strategies at the third location focused more on building traffic to an area and less on display optimisation.

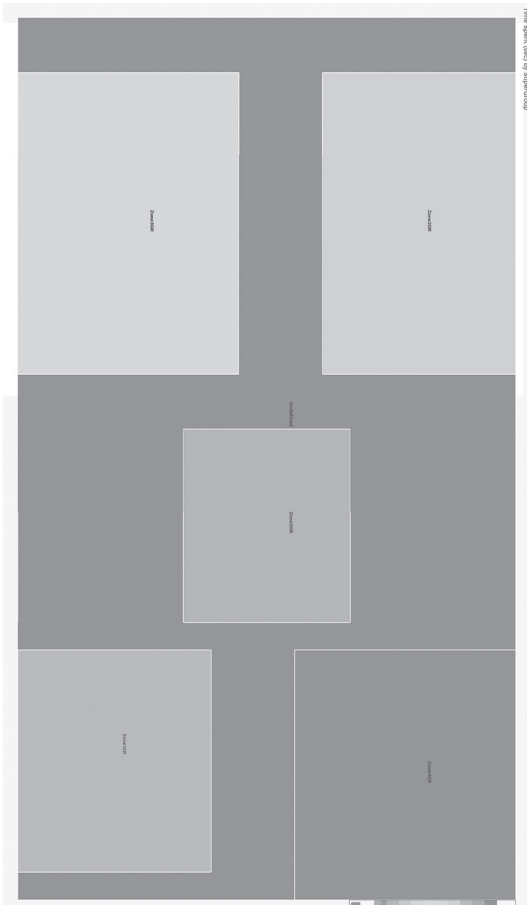

Figure 5: Shopper time by table area

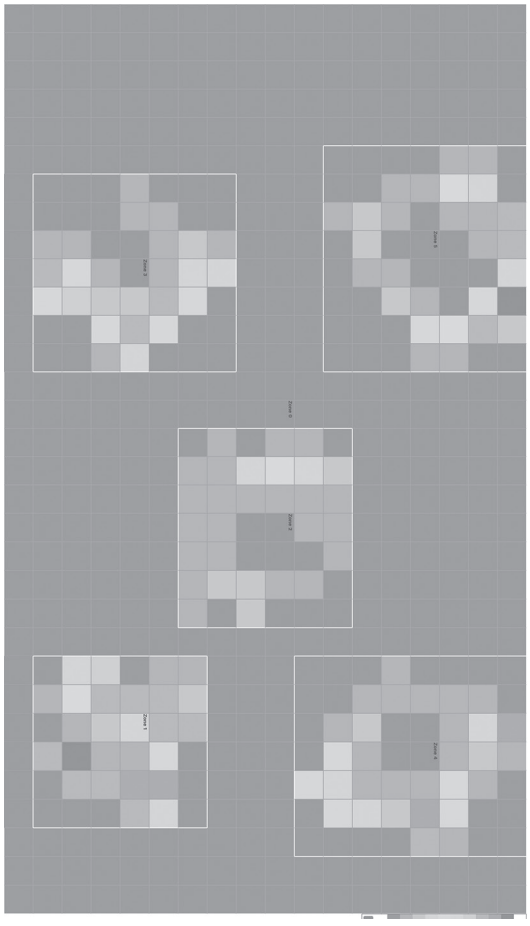

Figure 6: Product interactions by table location

For the two stores where traffic was not an issue, initial measurement suggested that roughly nine out of ten shoppers passed through or by the area without engaging with the displays. Of those who did stop, average linger times were short, but product engagement was pretty good.

Figure 3 shows what the flow pattern of traffic in the area looked like. The highest-traffic areas are in red, and the less trafficked areas are in blue.

It is easy to see that the main flow is along the along the right and left-hand sides and across the front face of the area (which, incidentally, leads to the cash-wrap).

Next, one can look at time in area (Figure 4).

Time, however, is a variable better studied at a more aggregate level. As part of the baseline, experiments were conducted with different views and aggregation types. Being able to build virtual store mappings in the planogram builder and then use them instantly makes it easy to iterate on lots of different ways of aggregating the store.

Figure 5 provides a view of three tables with specific areas carved out to study.

In addition to time, benchmarks were built around interaction, and product takeaways were developed for each display. With interactions, honing in on specific parts of the table display proved useful for driving a number of tests (see Figure 6).

This makes it possible to see whether specific faces or edges are important. Not only does this provide opportunities for optimising the product location and mix by area of the table, it also makes it possible

Table 1: Display area performance report

Description	Visits	Mean time spent	Linger rate (%)	Interaction rate (%)	Conversion rate (%)
Zone 1	17,385	6.2	14.2	10.30	2.2
Zone 2	25,199	5.1	18.1	6.20	1.6
Zone 3	16,001	5.0	14.4	7.40	1.5
Zone 4	13,287	6.4	14.9	9.50	1.4
Full area	45,365	11.4	38.4	13.10	1.8

to compare product effectiveness holding location in display constant. This is really important for tracking individual or brand-level performance in a display.

Seeing the visual flows is great, but testing requires hard numbers in order to track actual performance. With the grid-system used in building the digital maps, every metric is trackable at *any* level of store geography — from a square foot on up.

Table 1 presents a typical tabular funnel report that was used for tracking A/B tests.

Part 4: A Sample of the first test

With a baseline in hand, a series of tests was built into the project. Following is a little sample of what actual display measurement looks like and how the process works based on the very first test conducted:

- *Start with a five-table feature display area:* Figure 7 shows what the area looked like at the beginning of the project, with five tables — four boxing the area and one in the centre.
- *See how shoppers flow around displays:* As Figure 8 shows, the middle table was not in the flow of traffic. There were heavy flow areas around the area while the middle table was hardly noticed (Figure 9).
- *Test based on the data:* The first non-baseline test was to try moving the centre table adjacent to the natural traffic flow (Table 2). Moving it to the flow area made a staggering difference in performance across every metric.

This simple example highlights one of the main benefits of display analytics — the very straightforward path from analytics to action. Most good tests seem — in retrospect — obvious. And the more obvious a test seems, the more likely it is to produce measurable impact.

Part 5: Process

As the example highlights, the key to using these data it is to treat the data as a tool for building (and running) structured tests. It is nice to know how many shoppers passed a display and how many engaged with it. Furthermore, by comparing those numbers over time and across locations, one can get a general sense of how a store or campaign or season is working. That, however, is not really the best use of this kind of detailed data.

With initial tests demonstrating significant upside to changes in the display area, a gradual rollout of tests throughout the holiday season was conducted. Tests ran the gamut from simple product rotations designed to help understand whether product or location was driving interaction success to substantial disruptive display elements that dramatically altered the flow of traffic into the area. Most tests ran for at least two weeks, as this proved to be about the minimum amount of time to capture a good view of test results. Because store behaviours are dramatically different on weekdays versus weekends (as is total volume), running over multiple weekends at least two full sets of

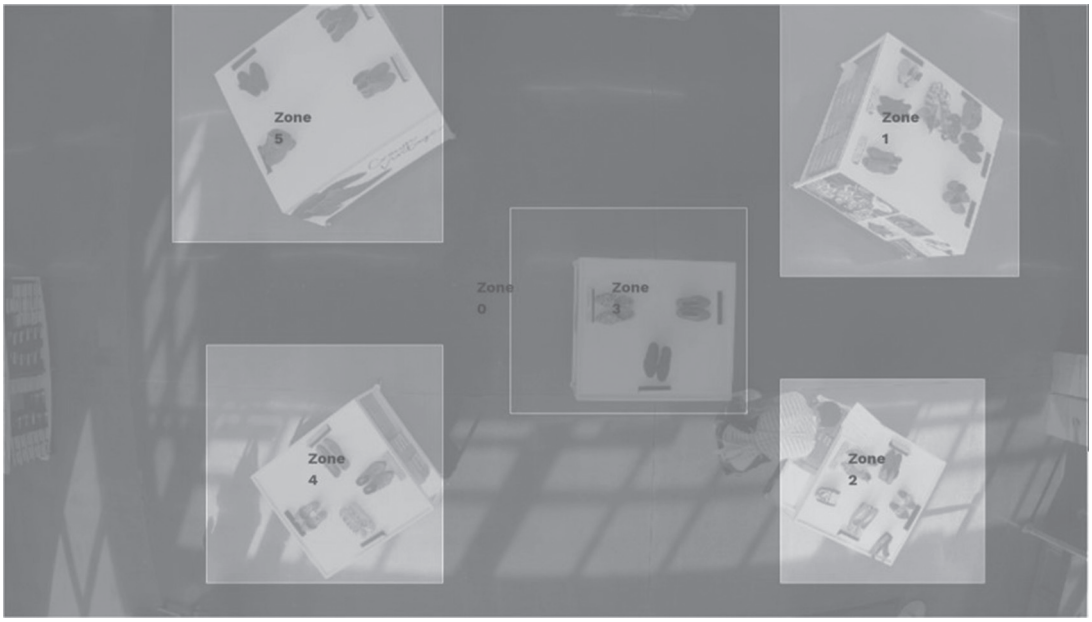

Figure 7: The basic feature area layout

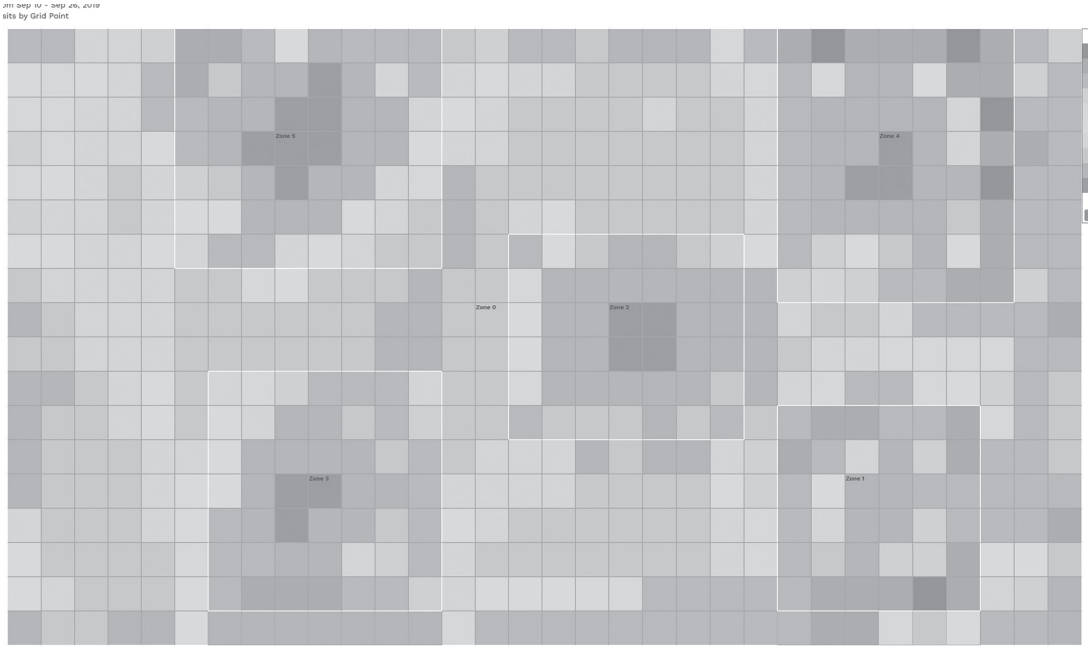

Figure 8: Initial flow of shoppers in the area

weekdays was necessary to collect statistically significant results.

Store display behaviours were also found to show significantly more day-to-day variability than is common in digital analytics. With high short-term variability, it is essential to let a test run long enough to provide stable and definitive results.

Among the more interesting tests were ones focused on optimising merchandise to match the flow in the area, adding a mixture of disruptive display elements

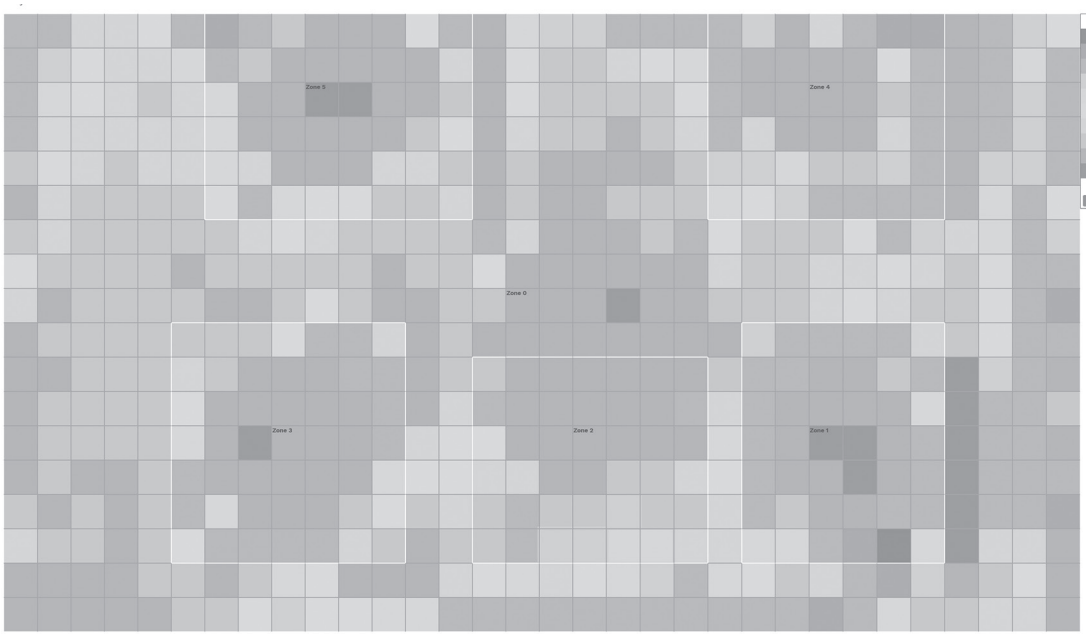

Figure 9: Moving the centre table adjacent to the natural traffic flow

Table 2: Sample test results reporting

Metric	Pre-test	Test	Improvement (%)
Visits per day	258	877	240
% of traffic	7	41	502
Interaction rate (%)	4	6	36
Product per day	12	54.0	363

that fundamentally redirected traffic and changing the product mix to create a single focus point. Even the angle of the tables turned out to be an interesting test point as it influenced both traffic flow and interaction hotspots.

Because the tests were run throughout the holidays, the analytics was also able to provide a definitive answer on just how feature area usage changes during that critical time of year.

What is particularly interesting is the ability to drive significant change without fundamentally re-engineering the visual display elements. Every test worked with existing visual display elements. But, of course, there is an unending panoply of tests that could be done by changing the visual display elements in terms of product, density, lighting and placement.

SUMMING UP

Experimentation is the single most important tactic in digital optimisation, yet it is hard to bring genuine experimentation methodologies to physical retail. Using advanced shopper measurement technologies, this no longer has to be the case.

Figure 10 shows what the whole process looked like.

The system not only answered all of the basic questions about usage and performance, it easily supported the kind of controlled experimentation that was desired.

The basic measurements included true opportunity, time spent, engagement percentage, interaction percentage and

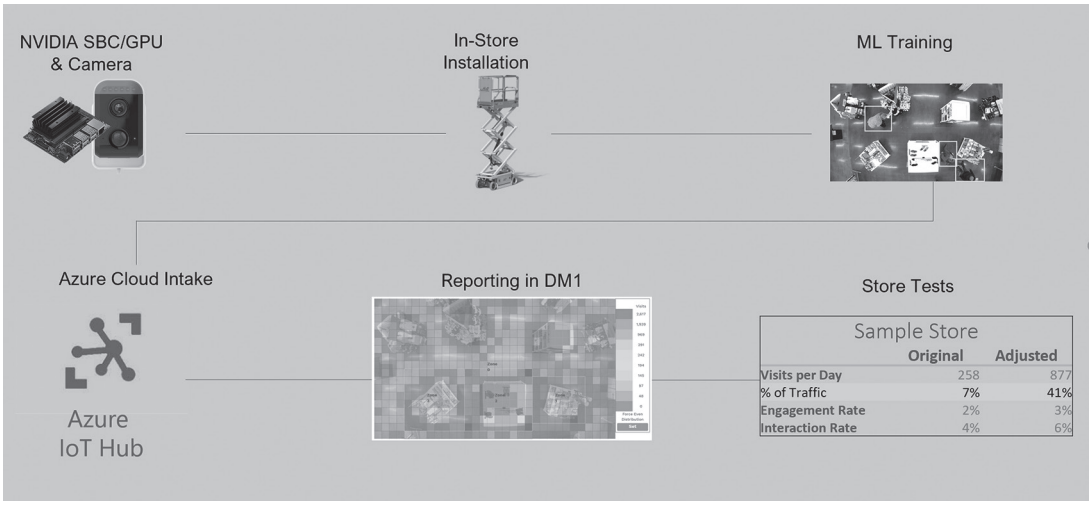

Figure 10: Display analytics project flow

product take-away rates. Going beyond rates, the flow data and interaction hotspots proved that the physical/geolocation aspects of the data can be put to rich use.

Best of all, the project proved that this is a rare area where test ideas are easy to generate and will often have significant performance impact. Digital marketers have long understood that one of the most challenging parts of building a successful test programme is ideation. Some user experiences are difficult to ideate on and expensive to change. Modern display is neither. Almost any multi-display area can use most or all of the tests outlined above and almost certainly end up with significant performance improvement. Furthermore, the opportunity exists for an almost endless series of potentially impactful tests.

Given the importance of visual merchandising to stores, display analytics measurement a prime target for retail analysts, store marketing and visual merchandising teams. It is a high-value focus point of the store that is small enough to make data collection and implementation straightforward and which makes the translation of analytics to action remarkably easy.