How we measure the impact of advertising on in-store sales, for one of our clients

A drive-to-store campaign, and a Time-Based Regression model applied to till data

Anonymised customer story - a sports retailer.

SUMMARY

The analysis establishes that advertising generated **+18% of incremental sales** on the advertised aisles.

High confidence level · 95% interval: [+3% | +32%]

A simple year-over-year comparison is somewhat naive: it hides a great many things, the store's own momentum from one year to the next, changes in the assortment, the retail calendar, and effects outside the campaign. Isolating the share that belongs to advertising calls for a data science approach.

That is what this kind of analysis makes possible. Once the foundation is in place, it serves to **arbitrate investments** and to answer the questions the store actually asks itself: is a 30% discount a winning move, should a channel's budget go up, is an operation worth what it costs.

Here is how we did it ↓

1 Findings and scope of the analysis

A retailer invests in Google and Meta advertising to bring shoppers into its store. The campaigns are geared toward store visits, and the ads showcase products from **a subset of aisles only**. The owner's question is simple: how many shoppers and how much revenue did this advertising bring that the store would not have made without it?

Answering that question calls for a point of comparison: what the store would have done without the campaign. Since there is a single point of sale, that point of comparison **is built over time**, from the store's own history. That is the model's job, and the subject of the next section.

The chosen indicator is the number of **shoppers**, that is, receipts containing at least one item from an advertised aisle. It is the most robust measure available: it withstands changes of brand and of price. Revenue is then derived from it, at the period's average basket.

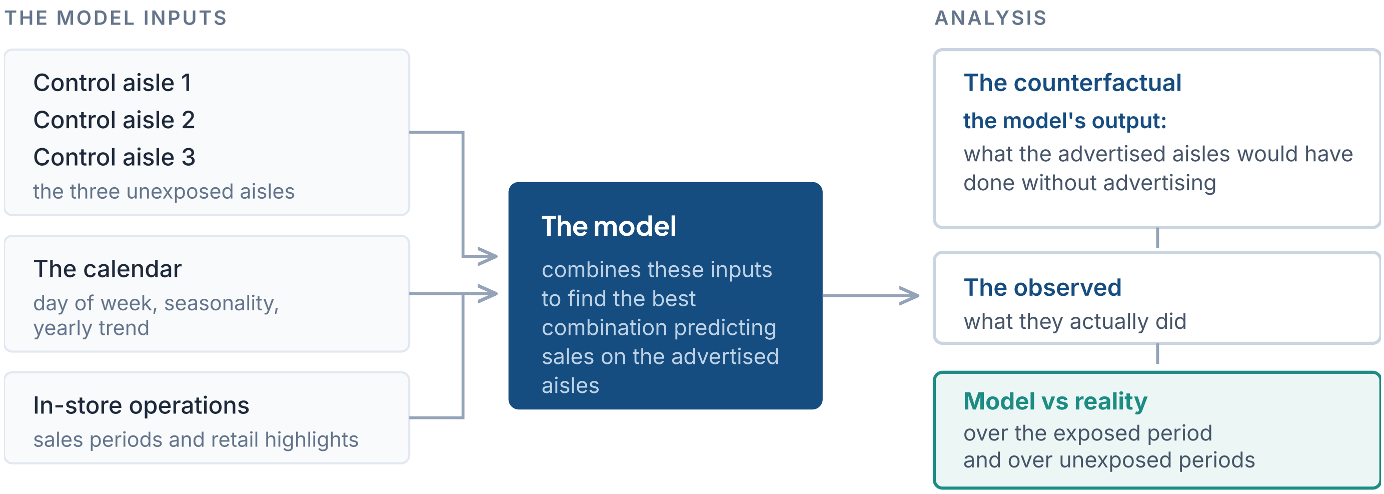
The main steps of the analysis

1. **Framing.** The question, the indicator and the measurement window, settled before the campaign.
2. **Data.** Daily sales history by aisle, enriched with the store's own context.
3. **Model.** Keeping the control aisles that err the least outside campaign periods.
4. **Validation, then measurement.** Checking the model's error on periods with no campaign, then reading the gap over the delivery window.

On the data side, the analysis needs **daily sales by aisle**, the store's **internal context** (promotion and sales-period calendar, retail highlights) and a history long enough for the model to learn the normal life of the point of sale: here, **18 months**.

2 The method

We build a **twin** of the exposed aisles from the store's aisles the campaign did not touch. Those aisles live the same life as the others: same weather, same footfall, same weekends, same season. The model captures those regularities from the till history and uses them to **predict** what the advertised aisles would have done without advertising. That is the counterfactual, and the measurement is then a subtraction: what actually happened at the till, minus what the model predicts.



The approach is a **Time-Based Regression**, a method published by Google in 2017 (Kerman, Wang and Vaver), designed for tests where only **few** control units are available: inference is read in time rather than in space. Adapted here on four points: one coefficient per control aisle rather than a single one on their sum, a logarithmic link rather than a linear relationship, the addition of the retail calendar, and control aisles chosen on their behaviour outside campaign periods rather than matched in advance.

The model is an equation learned from the data. These are not hand-picked coefficients: they are **estimated by computation**, by looking for the combination that best reproduces the opening days preceding the campaign.

The upstream framing work compared every possible combination of control aisles, and the trio retained **is not the most correlated** with the exposed aisles: it is the one that errs the least over periods with no campaign. An aisle that correlates strongly but carries its own momentum is a poor control, and correlation does not say so.

WHAT THE MODEL ABSORBS

The control aisles absorb everything that affects the store as a whole: **weather, mall footfall, public holidays, competitive pressure**. The calendar absorbs the weekly rhythm, the seasonality of sport and the store's year-over-year growth. The operations absorb sales periods and retail highlights. All of it leaves the comparison before anything is measured.

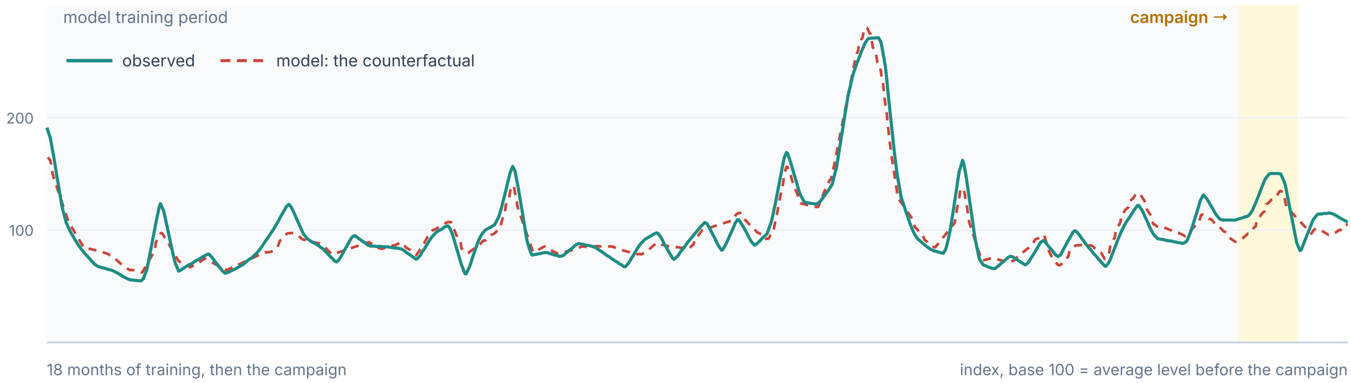
VALIDATING THE MODEL, AND ITS CONFIDENCE

The model is tested on **29 control periods** of the same length, taken before the campaign, where no advertising was running. Over those periods it finds **−2%**, a near-zero error: the model is close to reality. And none of those 29 periods reaches the level observed during the campaign.

That is what earns the right to read a gap during delivery. Without this step, a gap of a few percent cannot be told apart from the store's normal noise.

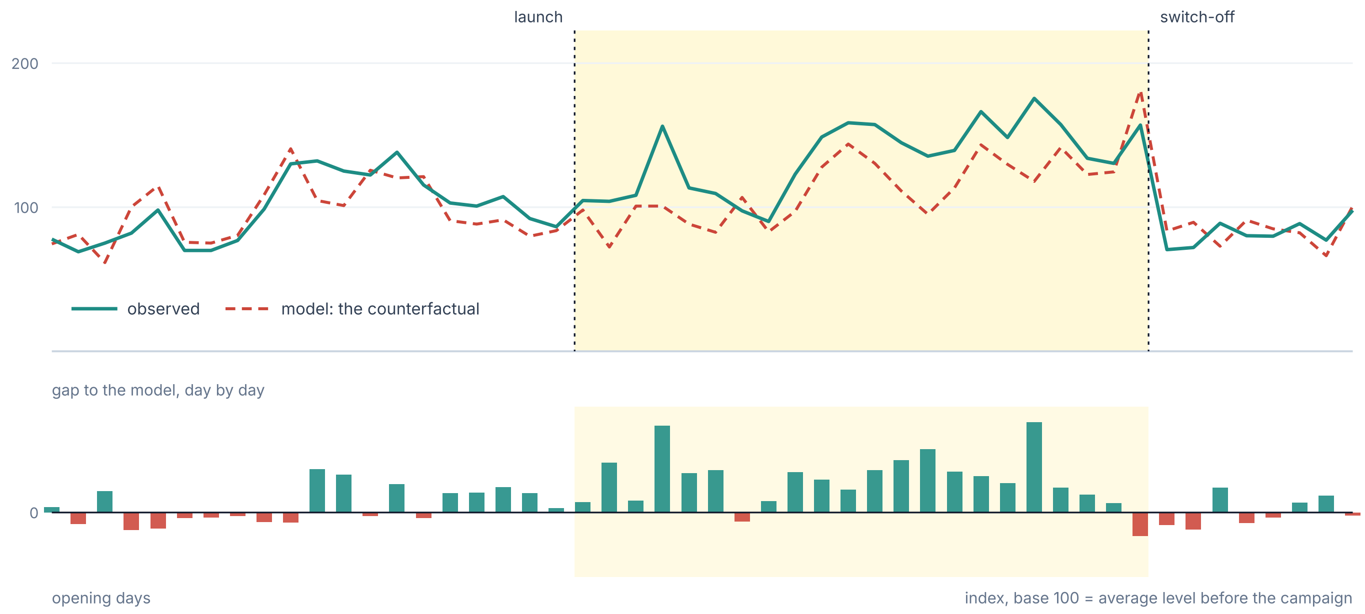
3 Analysing ad effectiveness

The model tracks reality across the whole available depth



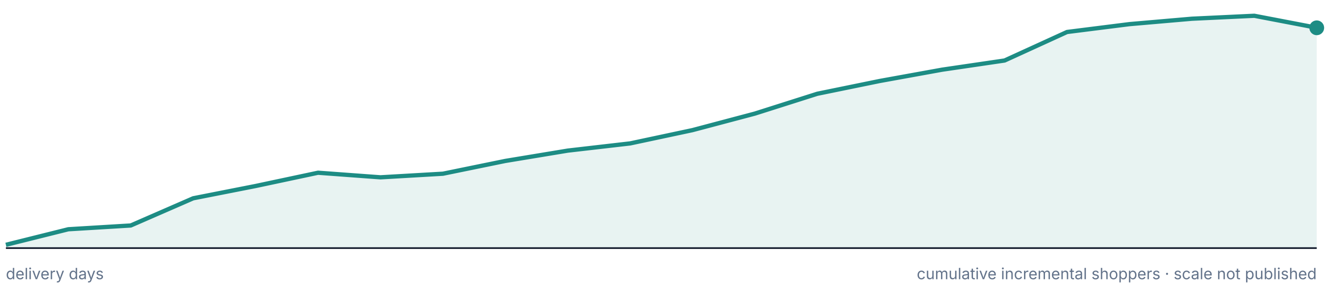
The model is well calibrated. It tracks observed activity across the 18 months of training, sales peaks included. Outside the campaign the two curves overlap; they separate clearly during it. Values are expressed as an index.

The gap opens at launch and closes at switch-off



The effect follows the delivery calendar. The gap opens at launch and closes as soon as the advertising stops. The lower panel shows it day by day: positive on **20 days out of 22**, negative on two. That is how a real effect behaves, not a smoothed curve.

Incremental shoppers, accumulated day after day



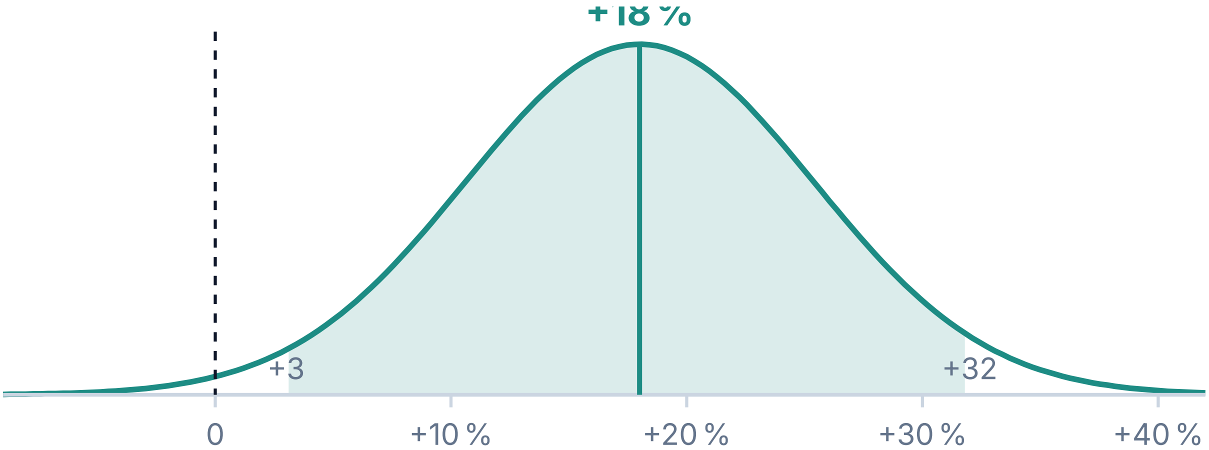
The volume builds up across the whole delivery period, without a break: every campaign day adds shoppers the model did not expect. The scale is not published, the case being anonymised; what matters is the progression.

THE RESULT OF THE MEASUREMENT

+18 %

incremental sales on the advertised aisles

95% range: **+3 % to +32 %**
High confidence level



The range sits entirely above zero. Illustrative shape: the interval comes from the dispersion measured over 29 control periods, not from an assumption of normality.

Performance signals are clear, around +18%, for an estimated ROAS above 4. That estimate includes the effects the measured scope leaves out: purchases made on other aisles by shoppers who came thanks to the advertising, and newly acquired customers who will return. On the strictly measured scope alone, the ratio is lower.

Two readings of the same sales

THE READING	WHAT IT ANNOUNCES
The year-over-year comparison same period, same set of aisles	**+40%**
The measurement what the model attributes to advertising, everything else neutralised	**+18%**

The **+40%** is real, but it cannot be used as it stands: it adds up the advertising, the store's natural growth, a deeper assortment than last year and the shift in the retail calendar between the two years. The method exists to sort through that. It yields a lower figure, one that keeps only what comes from the advertising.

What the ad platforms say

Alongside the model, platform data was reviewed: impressions, clicks, direction requests to the store and estimated visits. **They do not measure the same thing as the till**, and they capture only part of the real footfall: it is a low estimate. They point the same way as the measurement, with direction requests and estimated visits rising during delivery.

4 Conclusion

The advertising worked. Three independent readings converge: the model measures +18% of revenue on the advertised aisles, the effect follows the delivery calendar exactly, and platform signals confirm a strong intent to visit the store.

A DELIBERATELY CONSERVATIVE FIGURE

The +18% counts only the revenue made on the advertised aisles. A shopper who came for an advertised product often also buys an item from another aisle: all of that sits outside the measured scope.

Two uncounted effects point the same way: sales made on other aisles by shoppers who came thanks to the advertising, and newly acquired customers who will return. The real effect is therefore higher than the published figure, without the available data allowing us to quantify it.

The opportunity for the client

The analytical foundation is in place and directly reusable on upcoming campaigns. Three uses open up.

1. **Accelerating advertising investment** on measured grounds, rather than on intuition or on a year-over-year comparison.

2. **Arbitrating retail operations.** The same mechanism answers other questions: what a discount really brings, what a retail highlight produces, what one channel is worth against another.
3. **Making measurement a reflex.** The model is calibrated on this store: the next campaign is assessed without starting from scratch.

5 Notes and limits

NEUTRALISED BY CONSTRUCTION	NOT NEUTRALISED
Weather, mall footfall, competitive pressure, sales periods, the month's commercial mood. These factors affect the control aisles as much as the advertised ones: they leave the comparison.	**An outside event simultaneous with the launch.** Two events starting on the same day cannot be separated by any method. The figure covers the period as it presented itself.
Stock levels and assortment depth. The model learns on the months preceding the campaign: an improvement that happened before is already in its training. That is what a raw year-over-year comparison cannot do.	**The effect on the store's other aisles.** The available data does not allow us to settle it either way.

A complementary check covered stock levels. The exposed aisles carried noticeably more stock than the previous year, but that increase does not show up in sales outside the campaign. **Nothing therefore indicates that stock explains the measured effect**, and the campaign window stands out clearly from the rest of the year while stock is high all year round.

Choosing the method

The Time-Based Regression and modelling from control aisles are not the only route. That choice follows from a constraint: a single point of sale, hence no geographic unit to compare. On a network of several stores, we would have recommended a **geo-based approach**, where exposed areas are compared with areas deliberately held out of the campaign. Other methods could also have worked. **The method is chosen on the structure of the available data**, not by preference.

The specification

Poisson regression with a logarithmic link, trained on 18 months of daily sales. The Poisson distribution is the one for counts, and the logarithmic link makes effects readable as percentages whatever the level of the day. Inputs: three control aisles, the calendar (day of week, seasonality, trend) and the store's retail operations.

The interval

It comes from the model's measured behaviour, not from its assumptions: the full computation is replayed over 29 control periods taken before the campaign, and the dispersion is read off. The control aisles were chosen on one half of those periods, the confidence figures computed on the other half.

95% interval: the effect lies between +3% and +32%, for a central value of +18%. The analysis was also replayed with **CausalImpact**, on the same data and the same control aisles: two independent implementations, **the same conclusion to within 0.2 point**.

What remains uncertain

- The effect of an outside event simultaneous with the campaign launch, inseparable from its calendar.
- The effect on the unmeasured aisles, which the data does not allow us to settle.
- The seasonality of the end of the period, which rests on a single earlier occurrence.

Sources

The store's till software, daily sales by aisle over 18 months · Google Ads and Meta Ads platforms · retail calendar and stock statements provided by the store. Results are published as

percentages and indices, with no absolute values.

THE EDGEANGEL SERVICES THIS CASE DRAWS ON

Advanced Media Measurement

MMM, incrementality tests, causal measurement: establishing what each channel really contributes →

Modern Data Stack

The foundation that makes measurement possible: collecting, historising and securing sales data →

LET'S TALK

How much do your campaigns really bring in?

This measurement setup is built from your own sales data, for a single store as well as a network. A first conversation is enough to tell whether your data supports a solid measurement.

Discuss your project →