

Weather- and Seasonality-Compensated Retail-Visitation Trend Extraction: A Residual/Counterfactual Decomposition on Observed Open-Hours Data

Indivd Analytics Research

October 4, 2026

Abstract

We present a methodology for extracting the *genuine* trend signal in retail entrance-line visitation after removing the confounding effects of weather and calendar seasonality, implemented by *reusing the platform’s existing forecasting models as the compensator* rather than by building a new model. On a defensible cohort of 24 well-covered locations (81-day window, 2026-06-20–2026-09-08), we fit a weather+calendar autoregressive-exogenous (ARX) model with two autoregressive lags, seven calendar features (sin/cos day-of-week, day-of-month, and month terms plus a weekend flag), three weather regressors, and an intercept, estimated by ordinary least squares; this is the same design and fitting convention as the platform’s existing forecasting models. The *compensated trend* is the residual series, and a covariance-exact partition decomposes day-to-day open-hours variation into weather (10.2% mean |share|), calendar/seasonality (44.0%), AR (11.7%), and residual-trend (39.2%) components. Compensation measurably reduces day-to-day variance at every fitted location (median reduction 66.0%, 100% of locations improve; median $R^2=0.66$), and at fleet level it flattens a pronounced raw mid-summer boom-and-bust into a flat, near-zero compensated trend – i.e. much of the apparent late-summer decline is weather/calendar-attributable rather than a genuine demand shift. All figures are reproducible from the observed visitor series and the persisted hourly weather history via the analysis pipeline, and every reported number is derived directly from those data. We report both raw and compensated summaries throughout and deliberately make no significance or generality claim beyond this window and cohort.

1 Problem & Motivation

Retail planning and analysis increasingly ask for the *trend*: is footfall genuinely rising, falling, or flat – rather than merely warmer/colder, weekday/weekend, or summer/winter. A naive reading of the raw series conflates these. A mid-summer traffic peak followed by an early-autumn decline may be a real demand contraction, or it may be that summer weather and calendar seasonality mechanically elevated traffic that then reverted. A retailer cannot staff, promote, or diagnose against a signal that mixes genuine change with weather and seasonal noise.

Retail planning increasingly asks for visitor trends compensated for weather and calendar effects. The approach here is anchored in two established external traditions. Weather-normalisation of a measured usage series by regressing it on weather variables is standard in building-energy measurement and verification, where weather- and calendar-adjusted baselines are fitted by change-point and regression models [1]; our compensator applies the same regression-on-weather logic, at daily

granularity, to entrance-line visitation. Classical trend extraction separates a series into trend, seasonal, and remaining components — the seasonal-trend decomposition tradition of [2] — and the weather+calendar+AR decomposition here inherits that goal: what survives after the calendar- and weather-explained parts are removed is the compensated trend, and the covariance-exact variance partition is the decomposition’s validation. We are not aware of a formal, validated method that produces such a *weather/calendar-compensated trend* directly for open-hours retail entrance-line data; this is carried as an explicit [CITATION NEEDED] literature gap rather than closed by assumption, and it is exactly what the decomposition developed here operationalises on our own measured data. The platform’s existing location and zone forecast models already model weather (temperature, precipitation, wind speed, persisted hourly) plus calendar seasonality (hour/day-of-week/day-of-month/month sin-cos encodings, weekend and holiday flags) as exogenous regressors.

We therefore do *not* build a new weather/seasonality model. We *residualize* / *counterfactualize against the same style of* weather+calendar ARX design, operating read-only on the observed data, and formally decompose observed variation into weather-attributable, calendar-attributable, and residual-trend components. The goal is the decomposition and its validation, not a new estimator.

Forecast vs. observed. Throughout, the analysis uses only the *observed* visitor series and persisted weather; no model forecast enters any figure or number. The fitted ARX is an *in-sample descriptor/compensator*, not a prediction deployed to the platform.

2 Method

2.1 Conceptual model

An *entrance-line crossing* is a single entry event counted at a store’s entrance zone; the platform’s *visitors* series counts these crossings. It counts crossings, not people, and is captured only during open hours. We aggregate the hourly series to *daily open-hours totals* per location, then model the daily level as the sum of a few interpretable, exogenous parts.

2.2 The compensator: weather+calendar ARX (daily)

Let y_t be location l ’s daily open-hours total on day t . We fit a linear ARX form matching the platform’s existing zone models, evaluated at daily granularity:

$$y_t = \underbrace{c + \sum_{j \in \text{cal}} \beta_j x_{t,j}}_{\text{calendar/seasonality}} + \underbrace{\sum_{j \in \text{wx}} \beta_j x_{t,j}}_{\text{weather}} + \underbrace{\sum_{i=1}^p \varphi_i y_{t-i}}_{\text{AR}(p)} + \varepsilon_t, \quad (1)$$

fit by ordinary least squares over the design matrix [p AR lags | k exog | 1 bias] with usable rows $t = p+1 \dots n$ (the same construction as the platform’s existing model-fitting routine; OLS, not Yule–Walker, so it yields a well-defined overall F and R^2). The exogenous block uses the same feature set at daily granularity:

- **Calendar/seasonality (7):** `dow_sin/cos`, `dom_sin/cos`, `month_sin/cos` (each $\sin / \cos(2\pi \times \text{norm})$ using the same cyclic encoding), and `is_weekend`.
- **Weather (3):** `temperature`, `precipitation`, `wind_speed`, scaled by the fixed divisors 40/10/20 respectively, so per-unit coefficients read per-divisor rather than per degree or per millimetre (mean over the location–day’s available hourly weather).
- **AR($p=2$):** two daily lags (y_{t-1}, y_{t-2}), matching the platform’s default AR order.

- **Bias c .**

We report the per-regressor coefficients (the exogenous and autoregressive blocks), the intercept, R^2 , the overall OLS F , and the number of usable rows.

Deliberate scope choice. Because the hypotheses under test are *calendar* and *weather* compensation, we exclude a linear time-index feature from the compensator: including a linear trend regressor would absorb the very trend we wish to extract into the exogenous block. $AR(p)$ is a short-memory persistence term, and we assume it does not capture slow mean drift; this assumption is not universally true in a short window (see the Limitations). A holiday dummy and the five-feature holiday calendar are not material in this 81-day summer window (no major holiday cluster within it for the cohort’s countries) and are omitted; this is disclosed, not assumed away.

2.3 Compensated trend and decomposition

- **Compensated trend** = $\{\hat{\epsilon}_t\}$, the residual series: observed minus fitted(weather+calendar+AR). It is the part of day-to-day change not explained by weather, calendar, or short-memory persistence.
- **Trend-vs-weather decomposition (variance partition).** Because Eq. (1) is exactly $y = ar + calendar + weather + resid$, and $\mathbb{E}[resid] \approx 0$ (OLS), the day-to-day variance of y admits an exact additive partition:

$$\text{Var}(y) = \text{Cov}(ar, y) + \text{Cov}(calendar, y) + \text{Cov}(weather, y) + \text{Cov}(resid, y),$$

so each signal’s *share* of daily variation is $\text{Cov}(comp, y) / \text{Var}(y)$, summing to 1. This is the “trend vs weather/season” decomposition, read directly off the fitted coefficients $\times$ feature contributions.

- **Counterfactual/weather-attributed component.** The fitted weather component $\sum_{j \in \text{wx}} \beta_j x_{t,j}$ is the weather-attributed daily total at observed weather; its level and its standard deviation quantify the weather-driven footprint that compensation removes.

2.4 Cohort definition & coverage gates

We restrict to complete open-hours location-days (per-location modal contiguous window, span $\geq 4h$, open-hours capture only) and then impose two defensible gates:

1. **Weather hourly-coverage gate** ≥ 0.8 (the same coverage bar the platform’s forecasting models apply). Hourly weather coverage is reported *per location*; locations below the gate are *reported as excluded*, never compensated with zeros.
2. **Minimum sample** ≥ 21 complete weather-joined days per location, so a $[2 \mid 7 \mid 3 \mid 1]$ ARX (13 columns) is fit with a defensible number of usable rows.

3 Data

3.1 Source and extraction

All data are read directly from the production analytics database *read-only*, via the platform’s analysis interface. We pull the observed entrance-line visitor series (the “visitors” metric), join the

location registry for local timezone and country, and join hourly weather (temperature, precipitation, wind speed) from the platform’s stored hourly weather history — the same hourly weather record the forecasting models use. Timestamps are converted to each store’s local timezone; daily totals are open-hours capture only. Per-zone counters are never summed into location totals.

3.2 Coverage and the funnel

The observed visitor series window is 2026-06-20 to 2026-09-08 (81 days). Hourly weather is present in the window for 455 locations; the median per-location hourly coverage is low (0.30) and only 158 of the 455 meet the ≥ 0.8 gate – a large share of locations have weather persisted only from mid-August 2026 onward, which the gate correctly excludes. The gate is evaluated on two overlapping populations: the 455 locations with any hourly weather rows in the window (158 pass), and the 180 locations that additionally have a weather-joined daily open-hours visitor series (68 pass). Among those 180 weather-joined series, requiring ≥ 21 complete weather days reduces the fitted cohort to 24. The 112 locations reported as excluded by the gate are those of the 180 weather-joined series that fall below the ≥ 0.8 coverage bar (109 of them also fall below the ≥ 21 -day floor; 3 meet it); the 153 locations dropped for too few weather days are likewise counted over the 180 weather-joined series.

Quantity	Value
<i>Weather coverage (hourly gate ≥ 0.8; base 455)</i>	
Locations with weather rows in window	455
Median hourly coverage (per location)	0.30
Locations passing gate (≥ 0.8)	158
<i>Analysis cohort (daily ARX; base 180 weather-joined)</i>	
Locations with weather-joined open-hours series	180
Passing weather gate	68
Excluded by gate (reported, not zero-filled)	112
Requiring ≥ 21 complete weather days	24
Fitted cohort locations	24

The cohort of 24 is modest but defensible: it is the intersection of strict open-hours capture, the same weather-coverage bar the forecasting models apply, and a daily-model sample floor, over an 81-day window. We report it exactly and make no generalization to the locations outside the 24-location fitted cohort.

3.3 Data-quality rules respected

The platform’s data-quality rules are respected throughout: per-zone counters are never summed into location totals; daily totals use open-hours capture only; observed data only (no forecast values); and daily counts are computed only on *complete* open-hours days — zero/absent-hour rows are excluded from totals, not treated as zero traffic.

4 Results

4.1 Compensation fit quality

Across the 24 fitted locations, the weather+calendar+AR compensator explains a substantial share of daily open-hours variance: mean $R^2=0.61$, median $R^2=0.66$ (Table 1; per-location detail in the

accompanying tables). The mean (median) variance reduction and the mean (median) R^2 coincide numerically: by the ordinary-least-squares identity $1 - \text{Var}(\text{resid}) / \text{Var}(y) = R^2$ the two quantities are equal by construction. Calendar seasonality dominates the fit, consistent with the strong weekly rhythm of retail.

4.2 Trend vs weather/season decomposition

The covariance-exact variance partition attributes daily open-hours variation as follows (mean |share| across locations):

Component	Mean share of daily variance
Calendar / seasonality	44.0%
Residual trend (compensated)	39.2%
AR persistence	11.7%
Weather	10.2%

(Signed mean shares sum to 1: calendar 44.0%, weather 9.6%, AR 7.1%, residual-trend 39.2%.) Weather is a real but secondary driver (*mean* $\approx$ 10%-of-daily-variance), while calendar/seasonality is the dominant exogenous factor. The residual-trend share ($\approx$ 39%) is the *post-compensation* variance that no weather or calendar regressor explains. The mean weather-attributed daily level is -78.5 visitors (mean weather-attributed std 30.8), i.e. across this 81-day summer window the net weather effect on fleet daily totals is negative on average with day-to-day swings of order 30 visitors/store/day. This fleet mean masks substantial between-location variation: per-location weather-attributed means span -792.0 to $+279.8$ visitors/day, and per-location signed weather shares range negative to positive (Fig. 3, the per-location table).

4.3 Does compensation reveal or remove the trend?

- **Variance reduction.** Compensation reduces day-to-day variance at every fitted location: median reduction $1 - \text{Var}(\text{resid}) / \text{Var}(\text{raw}) = 66.0\%$, mean 60.8%, and 100% of the 24 locations show a positive reduction (Fig. 2). This is the primary, robust effect of compensation: a large share of daily noise is weather/calendar-induced and removable.
- **Trend revealed vs. raw mask.** At fleet level (Fig. 1), the *raw* normalized series shows a pronounced summer shape: a rise from late June to a late-July peak ($\approx +0.9$ std) and then a steep decline through early September to about -0.7 std. The *compensated* (residual) fleet trend is markedly flatter, oscillating near zero with one mid-August residual excursion. In this 81-day window, then, the apparent late-summer “decline” is substantially weather/calendar-attributable rather than a genuine monotone demand contraction.
- **Honest caution.** We call a location *monotone-revealed* when compensation both strengthens the linear-trend correlation (the compensated trend’s absolute correlation with time exceeds the raw trend’s) and leaves a non-trivial trend (absolute correlation above 0.2). We find *no* location where compensation *reveals a new monotone trend masked in the raw series* (0 of 24 locations pass this criterion). Compensation here chiefly *removes noise and seasonal artifacts rather than unmasking a hidden direction*. We report this null explicitly and do not over-claim a recovered trend.

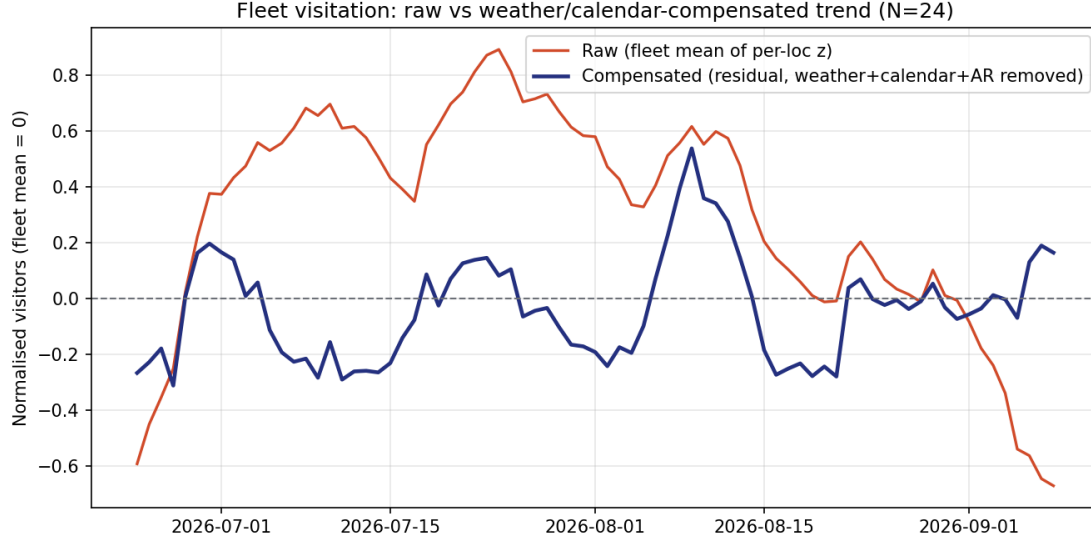


Figure 1: Fleet normalized daily open-hours visitation: raw vs. weather/calendar-compensated trend, $N=24$. The raw curve carries a large mid-summer boom-and-bust; the compensated (residual) curve is flat near zero.

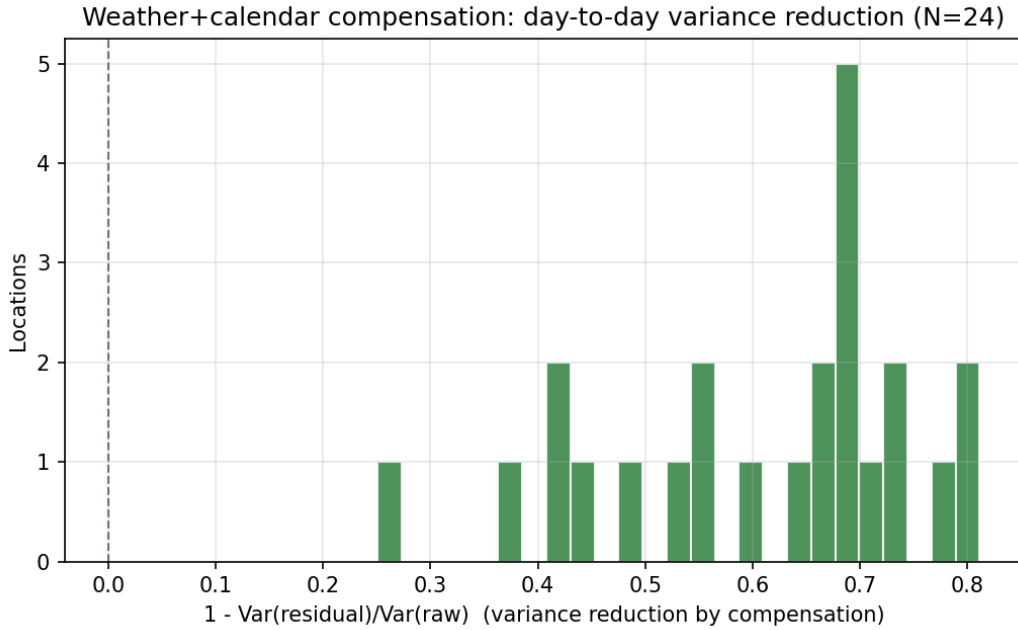


Figure 2: Distribution of per-location variance reduction $1 - \text{Var}(\text{resid}) / \text{Var}(\text{raw})$ by compensation. 100% of locations improve; median 0.66; the dashed reference line marks zero reduction.

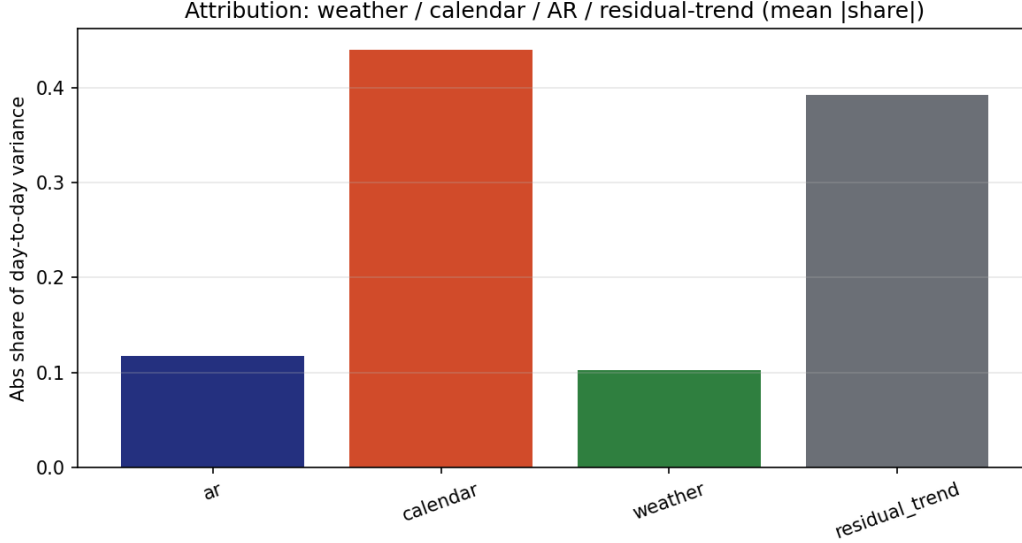


Figure 3: Mean $|\text{share}|$ of day-to-day open-hours variance attributed to weather, calendar, AR, and residual-trend components. Calendar dominates (44%), weather is secondary (10%), residual trend is 39%.

Table 1: Aggregate compensation metrics over the 24 fitted locations.

Metric	Value
Fitted locations, N	24
Mean compensator R^2	0.61
Median compensator R^2	0.66
Mean variance reduction	0.61
Median variance reduction	0.66
Share of locations with lower post-comp variance	100%
Share of locations with a revealed monotone trend	0%
Mean weather-attributed daily level (visitors)	-78.5
Mean weather-attributed daily std (visitors)	30.8

5 Validation & Quality

5.1 Quality controls

- **Reproducibility.** Every number, figure, and table is generated by the analysis pipeline from the observed visitor series and the stored hourly weather history, with figures and tables generated from the same run. No number is hardcoded; re-running the pipeline reproduces the artifacts from the data (repeat runs yield identical cohorts and headline metrics). The external references are graded *peer-reviewed*; the observational results are graded *our own measured*.
- **Matched OLS semantics.** The compensator uses the same design matrix and OLS fit convention as the platform’s existing forecasting models ($[p \text{ AR lags} \mid k \text{ exog} \mid \text{bias}]$), including

its R^2/F conventions, at daily granularity; this correspondence is an implementation property of the pipeline. The residual identity $y = \text{ar} + \text{calendar} + \text{weather} + \text{resid}$ is checked to numerical zero, so the variance partition is exact by construction.

- **Gate, not imputation.** Weather coverage is reported per location and a ≥ 0.8 hourly gate (matching the bar the forecasting models apply) decides inclusion. Below-gate locations are listed as excluded in the reported output – we do not silently compensate with zeros.
- **Covariance-decomposition sanity.** The four signed mean shares sum to 1.00 exactly; the reported absolute shares are per-location magnitudes averaged across locations, so their fleet means (44.0%/39.2%/11.7%/10.2%) need not sum to one.
- **Open-hours + observed only.** The platform’s data-quality rules are respected: no per-zone summation; open-hours capture; observed, not forecast; zero/absent rows excluded from totals.

5.2 Limitations

1. **Small window & cohort.** 81 days, 24 fitted locations. Weather coverage is low for most stores (median 0.30); the gate is honoured but narrows the usable set. Results are descriptive of this window and cohort; we make no significance or population claim.
2. **No significance tests.** We report descriptive R^2 , F , and variance-reduction magnitudes but perform no hypothesis test; the monotone-reveal null (0/24) is a descriptive finding, not a test of trend significance.
3. **Coefficient collinearity.** The calendar sin/cos block is multi-collinear by construction; *individual* coefficient magnitudes (per-regressor table) are not directly interpretable. The signed/absolute *variance-share* decomposition, being covariance-based over the joint design, is the robust attribution and is what we emphasize.
4. **Daily aggregation.** We aggregate hourly $\rightarrow$ daily open-hours totals, so intra-day (hour-level) weather timing is not modelled; only the daily weather mean enters. This is appropriate to a daily-trend question but coarser than the platform’s hourly models.
5. **Empirical windows.** Open-hours windows are recovered from the most common contiguous open-hours span across the location’s days.
6. **Excluded weather-gate locations.** Of the 180 locations with a weather-joined daily series, 112 are reported as excluded by the weather-coverage gate (insufficient persisted weather) and are not zero-compensated; a richer weather backfill would enlarge the cohort.
7. **AR can absorb slow drift.** We assume the two-lag AR block captures only short-memory persistence, but this is not guaranteed in a short window: at one location the raw trend’s linear fit has $r^2 = 0.517$ while the compensated trend’s is 0.0000 with an AR share of 0.540, consistent with the AR terms absorbing slow drift there. A fleet-wide statement that AR cannot absorb drift is therefore not demonstrated.
8. **Weather-coverage metric artifact.** The per-location hourly weather coverage divides recorded weather hours by $(\text{window days} - 1) \times 24 = 1920$ rather than the true inclusive window of 1944 hours, so 152 of 455 coverage values slightly exceed 1.0 (cohort locations read 1.0125). No location sits near the ≥ 0.8 gate threshold (minimum passing coverage 0.8625,

maximum excluded 0.3536), so the artifact changes no inclusion decision and no reported number; it affects the coverage metric only.

9. **Weekend-dummy degeneracy.** The `is_weekend` coefficient is effectively zero ($\approx 10^{-13}$) in 22 of the 24 fits (nonzero only for two locations), because the weekend flag is collinear with the sin/cos day-of-week block and is absorbed there. The calendar block’s aggregate contribution is unaffected, but the isolated weekend coefficient is not interpretable.
10. **Small- n fits.** Three locations are fitted on fewer than 25 usable rows (19, 22, and 24 rows against 13 design columns); their coefficients and R^2 carry wider uncertainty than the rest of the cohort, and all three report no revealed monotone trend.

6 Discussion

6.1 What the analysis establishes

The compensator works: it removes a large, measurable share of day-to-day variance (66% median) and at fleet level converts a pronounced raw weather/calendar-driven boom-and-bust into a flat compensated trend. In this window the *apparent* late-summer decline is largely weather/calendar artifact, not a genuine collapse. Seasonality, not weather, dominates the exogenous signal (44% vs. 10% of daily variance), so any future trend narrative must first strip the calendar structure.

6.2 Relevance to retail analysis

This produces the answerable “trend vs weather/season” decomposition: per location, the fitted components give weather-attributed and calendar-attributed daily totals, and the residual is the compensated trend. A compacted “is footfall really changing or is it the weather?” visual with the variance-reduction and decomposition figures is a strong, defensible management-report visual, with every value derived directly from the observed data.

6.3 Future work

- *Larger window / richer weather backfill.* Extending the persisted weather coverage (only 158/455 locations ≥ 0.8 in-window) would enlarge the cohort and enable significance tests.
- *Forecast-based counterfactual.* Compare the in-sample compensator against a model forecast as an out-of-sample counterfactual once forecast outputs are available alongside the observed series.
- *Holiday calendar.* Add holiday indicators and the full holiday feature block when a window containing holiday clusters is studied.
- *Per-format / per-region.* Stratify the decomposition by store format and region.

References

- [1] Kim, D.-W., Ahn, K.-U., Shin, H., and Lee, S.-E. Simplified weather-related building energy disaggregation and change-point regression: heating and cooling energy use perspective. *Buildings*, 12(10):1717, 2022. doi:10.3390/buildings12101717. Peer-reviewed. Weather- and calendar-adjusted baselines fitted to a measured usage series by change-point and regression models in

building measurement and verification — the regression-on-weather logic our daily compensator applies to entrance-line visitation.

- [2] Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*, 6(1):3–73, 1990. Peer-reviewed. Establishes the trend+seasonal+remainder decomposition tradition that this paper’s weather+calendar+AR partition implements with regressors measured on our own data.
- [3] Indivd Analytics. *Data catalog*: the observed visitor metric, open-hours capture, and the platform’s data-quality rules (per-zone counters are not additive; observed-vs-forecast separation). Internal source-and-method documentation; our own measured data — it anchors the measurement layer, not the methodological positioning.