

# An Anonymized Bounded Event-Contrast Study of Daily Retail Footfall around a Major International Sporting Event

Indivd Analytics

September 27, 2026

## Abstract

We report a bounded, descriptive contrast of daily retail footfall at a single duty-free retail outlet serving international transit passengers at a major aviation hub (hereafter **Site B**), comparing a major international sporting-event window against immediately adjacent pre- and post-window phases. Footfall is measured at the day granularity as the sum over the outlet’s 17 passenger zones, each zone-day observation derived from an authoritative day-grain visitation series. The event window spans 2026-06-11–2026-07-19 (39 days); the pre-window baseline is a 41-day season-adjacent shoulder (2026-05-01–2026-06-10), widened from an initially planned 10-day window after a data-coverage check showed the daily series reaches further back; the post window runs 2026-07-20–2026-09-25 (68 days). The store-aggregate daily series was 100% complete over the window across all 17 zones (no flat feeds, no missing calendar days), clearing the pre-specified data-sufficiency requirement for proceeding. On the raw medians, the event-window daily median was 1.5 % above the pre-window median but the raw mean was -1.1 % below it. A day-of-week-adjusted phase regression (phase indicators plus day-of-week fixed effects, cluster-robust by calendar day) estimates a -2170 visitors/day event-window contrast versus the pre baseline (95% CI [-7895, 3556],  $p = 0.458$ , not significant), and a randomization test that re-estimates the contrast under 1,486 random event-window boundary placements gives a two-sided tail probability of  $P = 0.843$ : the observed contrast is entirely typical of random phase placement. The event-window contrast is *not* robust to the pre-baseline specification: excluding the annual religious-holiday peak that falls inside the widened pre-window raises the estimate to +3.2 % median (DOW-adjusted 5013 visitors/day, 95% CI [854, 9171],  $p = 0.018$ ); the original 10-day pre-window read was +5.9 %. After the event window, footfall reverts to slightly below the pre-window level (median -7.0 % versus the pre median, DOW-adjusted  $p < 10^{-4}$ ). We report the full pre-baseline-robustness battery, the confound disclosure, and the data-quality and reproducibility checks; the contrast is honestly labelled as descriptive and is not a causal estimate of the event’s effect.

## 1 Problem and Motivation

Major international sporting events attract large mobilized populations and are frequently assumed to lift retail and airport commerce; the hosting and transit geographies around such events draw public and commercial interest. Estimating a causal effect of an event on a single retail outlet is, however, confounded by seasonality, passenger-mix changes, and the absence of a plausible counterfactual store. Rather than claim a causal league, we pose a narrow, answerable descriptive question: *did daily footfall at a single airport transit-duty-free outlet rise during the event window relative to an immediately adjacent pre-window, measured at the day granularity and adjusted for day-of-week, and is any such contrast robust to the choice of the pre baseline?*

This question is worth answering because it is directly answerable from the data we can measure at Site B, and because an honest, bounded answer – including a null or non-robust one – is more informative than an inflated headline. The null result and the pre-baseline sensitivity we document are the primary findings, consistent with a growing literature that questions oversized event-impact claims [6, 7].

### 1.1 Scope, terminology, and anonymity

We study **Site B**, a single duty-free retail outlet serving international transit passengers at a major aviation hub. To keep this work publication-ready and free of identifying information, we (i) refer to the

outlet only as “Site B”; (ii) omit all venue, city, airport, and operator identifiers; (iii) keep all plotted volumes *relative* (indexed to the pre-window median), so that figure labels never carry a raw magnitude that could be linked back to the outlet — absolute daily visitor counts appear only in the text and tables; and (iv) label the outlet’s internal passenger zones only as  $Z_1, \dots, Z_{17}$ . The exact number of zones (17), the event window, and the phase boundaries are stated, as they are not identifying on their own. The underlying event is a globally known international sporting event; the hosting geography is deliberately not named.

## 2 Method

### 2.1 Bounded window-contrast design

We partition the study window 2026-05-01–2026-09-25 into three contiguous phases defined by the event window:

- *Pre-window* (2026-05-01–2026-06-10, 41 days): an immediately season-adjacent shoulder. The study initially planned a 10-day pre-window; a data-coverage check showed the authoritative daily series extends several months further back, so the baseline was widened to a 41-day, 100%-complete window covering the full season-adjacent shoulder. This lengthening is itself material: the appended days contain a substantial annual religious-holiday peak, which we isolate in the sensitivity analysis (Section 5).
- *Event window* (2026-06-11–2026-07-19, 39 days; hereafter WC).
- *Post window* (2026-07-20–2026-09-25, 68 days): the reversion check immediately following the event.

The contrast is a within-outlet, phase-level descriptive difference. It is explicitly *not* a difference-in-differences design: there is no cross-sectional control, and no claim of a causal effect is made [5]. The study scope (data window, outlet, grain) was fixed in writing before the analysis, and its single change — the pre-window widening — is recorded as a dated amendment to that written scope record.

### 2.2 Day-of-week-adjusted estimator

The primary estimator is a day-level linear regression of daily total footfall on phase indicators with day-of-week fixed effects:

$$Y_d = \alpha + \sum_{k \neq \text{Monday}} \gamma_k \mathbf{1}[\text{dow}(d) = k] + \beta_{\text{WC}} \mathbf{1}[d \in \text{WC}] + \beta_{\text{post}} \mathbf{1}[d \in \text{post}] + \varepsilon_d, \quad (1)$$

where  $Y_d$  is the store-aggregate daily visitors at Site B, Monday is the reference day-of-week, and  $\varepsilon_d$  is clustered-robust by calendar day (single observation per day, so a Liang–Zeger cluster-robust sandwich [1]). The coefficient of interest,  $\beta_{\text{WC}}$ , is the day-of-week-adjusted mean daily contrast of the event window relative to the pre-window baseline; 95% confidence intervals and  $p$ -values are cluster-robust, following the guidance on inference with clustered errors [2, 3]. With a single observation per cluster, this inference is heteroskedasticity-robust but does not address serial correlation across days, a caveat we return to in the confound block. We report the store-aggregate (single daily observation per day) and a zone-pooled variant that pools the 17 zones with zone and day-of-week fixed effects (pre and Monday references), cluster-robust by calendar day over zone-day observations.

### 2.3 Placebo / permutation null

To characterize the null distribution of  $\hat{\beta}_{\text{WC}}$  under random *phase* placement, we draw 1,486 placebo event-window placements, each with a random start boundary that keeps each leg at least 15 days, re-estimate the day-of-week-adjusted contrast from equation (1) on each placebo placement, and take the two-sided tail probability  $P = \Pr(|\hat{\beta}_{\text{WC}}^{\text{placebo}}| \geq |\hat{\beta}_{\text{WC}}^{\text{observed}}|)$  as the randomization  $p$ -value [4]. The placebo null mean and spread are reported.

## 2.4 Minimum detectable effect

We report the minimum detectable effect (MDE) at  $\alpha = 0.05$  for the WC-versus-pre contrast,  $\text{MDE} = z_{1-\alpha/2} \cdot \text{SE}(\hat{\beta}_{\text{WC}})$ , using the cluster-robust standard error of the day-of-week-adjusted contrast coefficient; it quantifies the smallest WC-vs-pre contrast the design could detect at the observed noise level (the 80%-power variant adds the  $z_{0.8}$  term).

## 3 Data

### 3.1 Grain, source, and window

Footfall is measured at *zone-day* grain: for each of the outlet’s 17 passenger zones, the authoritative day-grain visitation series (day-level interval, local time–timezone suppressed for anonymity) is summed to a daily zone total and the store-aggregate is the sum over the zone-day observations (2,516 zone-day rows in-window; 148 store-days at 100% completeness). The analysis reads a retained, frozen snapshot of the production fetch (as of 2026-09-26); the entire pipeline is read-only, and the shipped numbers were re-verified against this frozen snapshot after manuscript generation. The shipped window is 2026-05-01..2026-09-25, recorded in a written scope document before the analysis ran.

### 3.2 Completeness and data-quality gates

Before estimating anything, the analysis applies a fixed data-quality battery: the day series is checked for flat feeds, all-zero days, and missing calendar days per zone, and the store-aggregate daily completeness is computed per day and over the event window. All 17 zones are 100% complete over the 39-day event window with zero missing calendar days and zero all-zero days (no flat feeds) over the full 148-day study window; the mean daily completeness over the event window is 1.0. The pre-specified rule was to abandon the study if mean daily completeness fell below 0.5 (more than half of visits unrecorded); the observed completeness clears it with wide margin, so the design proceeds on complete data.

## 4 Results

### 4.1 Phase descriptives

Table 1 and Figure 1 summarize the phase-level daily totals at Site B, indexed to the pre-window daily median (= 100). The event-window median sits at a modest 101.5 index versus the pre median (a 1.5 % raw-median difference), while the event-window mean falls -1.1 % short of the pre mean, reflecting asymmetric, holiday-inflated pre days. The post window runs -7.0 % below the pre median.

### 4.2 Day-of-week-adjusted contrast

Table 2 reports the primary regression results. The store-aggregate day-of-week-adjusted event-window contrast is -2170 visitors/day (95% CI [-7895, 3556],  $p = 0.458$ ): *not statistically significant*. The zone-pooled estimate is -128 visitors per zone-day ( $p = 0.447$ , similarly not significant). The post-window contrast relative to the pre baseline is -12019 visitors/day ( $p < 10^{-4}$ ), a significant reversion. The randomization test gives  $P = 0.843$  (null mean -6372, s.d. 5297): the observed contrast is entirely typical of a randomly placed window, independently confirming the non-significant regression result. The contrast-level MDE is 5725 visitors/day (4.14 % of the pre-mean), above the observed effect magnitude (Table 5).

### 4.3 Within-event temporal pattern

Splitting the event window into early, middle, and late thirds (Figure 2, Table 3), the middle third shows a *positive* median contrast of 10.3 % versus the pre median, while both the early (-0.6 %) and late (-1.5 %) thirds are at or below pre levels. This mid-window hump, coupled with the post reversion, is consistent with a transitory, window-bounded visitation burst rather than a uniform elevated plateau across the whole event window.

## 5 Sensitivity: pre-baseline robustness

Figure 3 and Table 4 report the central robustness finding. The widened 41-day pre-window includes an annual religious-holiday peak in its appended days; that peak inflates the pre baseline and *understates* any event-window contrast. Excluding the peak days from the pre-window raises the store median contrast to  $+3.2\%$  (DOW-adjusted 5013 visitors/day, 95% CI [854, 9171],  $p = 0.018$ ), and the original 10-day pre-window (which sits just after the peak) gives a  $+5.9\%$  median read. The same holiday exclusion changes the post-window read symmetrically (DOW-adjusted -4769 visitors/day,  $p = 0.009$ ), so the post reversion is not an artifact of the holiday either; both reads are reported rather than selected. The event-window contrast is therefore *not robust to the choice of the pre baseline* once the outlet’s seasonal and holiday calendar is held in-frame: the null, a modest positive, and a larger positive all obtain depending on how the pre is defined. This is the headline honesty result, and it bounds any claim the outlet’s own data can support.

## 6 Validation and Quality

### 6.1 Reproducibility and claim-verification

Every number in Sections 4–5 is produced by code reading one frozen results file, `results/output.json`; no figure, table, or prose line carries a hand-typed value. Re-running the analysis and regeneration scripts on the retained data reproduces every figure, table, and inline number exactly, and an independent re-run of the analysis from a fresh copy of the code and data reproduces the results file byte-for-byte. We ship this pipeline discipline with the paper so that any reader with access to the recorded data can regenerate every number from scratch; the manuscript regeneration hash for this build is 576b0e03cc1d.

### 6.2 Confounds and honesty block

We state the confounds that bound the interpretation:

- *Seasonality.* The outlet’s host geography experiences its annual summer peak in exactly the event window; the observed levels are thus *not* causally attributable to the event, and the seasonal growth into summer partly inflates any window contrast. The pre-baseline choice is the dominant lever (Section 5).
- *Passenger mix.* The outlet serves transit passengers; passenger mix (transit vs. origin/destination) is unmeasured, and a dedicated post-2022 passenger-flow source for the hub was not found in public data. The outlet is passenger-driven, not street-retail.
- *No counterfactual.* A single outlet, single event: there is no cross-sectional control, so the contrast is a within-store window contrast, not a difference-in-differences.
- *Window-boundedness and persistence.* Under the reversion read, footfall returns to (slightly below) pre-window levels after the window; residual elevation past the final peak days is flagged rather than hidden. The significant post-window drop is consistent with the elevation being concentrated in the event window, though a post-window seasonal decline is an equally plausible reading.
- *Granularity.* The day grain is authoritative for this series; an hour-grain defect class observed elsewhere in the wider estate is deliberately not used here.
- *Pre-baseline choice (primary honesty point).* The widened pre-window includes an annual holiday peak; the initial 10-day pre sits just after it. The two baselines produce materially different contrasts (Section 5), so the answer to our motivating question is *specification-dependent*, and we say so explicitly.

### 6.3 Limitations

The design is descriptive and single-store; power to detect a small lift is limited and the MDE (5725 visitors/day) exceeds the observed effect. Seasonality and passenger-mix cannot be separated from a genuine event effect with a single series. No causal claim is made.

| phase                 | days | median idx | mean idx | median $\Delta$ vs pre (%) |
|-----------------------|------|------------|----------|----------------------------|
| Pre                   | 41   | 100.0      | 103.8    | +0.0                       |
| Event                 | 39   | 101.5      | 102.7    | +1.5                       |
| Post                  | 68   | 93.0       | 95.1     | -7.0                       |
| event vs pre (median) | –    | 101.5      | –        | +1.5                       |
| event vs pre (mean)   | –    | –          | 102.7    | -1.1                       |

Table 1: Site B phase-level daily descriptives, indexed to the pre-window daily median (= 100); raw-median/mean event-window % contrasts versus pre; and post-window reversion. Generated from `results/output.json` by `code/gen_manuscript.py`.

## 7 Discussion

At the day-of-week-adjusted level and against a season-adjacent 41-day pre-window, daily footfall at Site B during the event window shows *no statistically significant* uplift (-2170 visitors/day,  $p = 0.458$ ), and a randomization test over random window placements does not reject the null either ( $P = 0.843$ ). The result is dominated by the pre-baseline construction: the annual holiday peak inside the widened pre suppresses the contrast, and removing it yields a modest, significant  $+3.2\%$  read ( $p = 0.018$ ) while the original 10-day pre yields  $+5.9\%$ . The event-window contrast is therefore not robust. The clearest within-window structure is a mid-window hump with post-window reversion, consistent with a transitory, window-bounded visitation pattern that reverts once the event ends. These findings align with evidence that large-scale event impacts on host and transit economies are often smaller, more local, and more temporally concentrated than headline estimates suggest [6, 7], and they underscore the value of reporting the pre-baseline-robustness battery rather than a single headline number. For a single-outlet study on a recently instrumented estate, a bounded, honestly labelled descriptive contrast – including its null and its non-robustness – is the defensible deliverable, and a causal estimate of the event’s effect on Site B is out of reach without a cross-sectional control.

## References

- [1] K.-Y. Liang and S. L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, vol. 73, no. 1, pp. 13–22, 1986. doi:10.1093/biomet/73.1.13.
- [2] A. C. Cameron, J. B. Gelbach, and D. L. Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, vol. 90, no. 3, pp. 414–427, 2008. doi:10.1162/rest.90.3.414.
- [3] M. Bertrand, E. Duflo, and S. Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, vol. 119, no. 1, pp. 249–275, 2004. doi:10.1162/003355304772839588.
- [4] M. D. Ernst. Permutation methods: a basis for exact inference. *Statistical Science*, vol. 19, no. 4, 2004. doi:10.1214/088342304000000396.
- [5] A. Abadie. Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, vol. 72, no. 1, pp. 1–19, 2005. doi:10.1111/0034-6527.00321.
- [6] A. Feddersen and W. Maennig. Sectoral labour market effects of the 2006 FIFA World Cup. *Labour Economics*, vol. 19, no. 6, pp. 860–869, 2012. doi:10.1016/j.labeco.2012.07.006.
- [7] C.-K. Lee and T. Taylor. Critical reflections on the economic impact assessment of a mega-event: the case of 2002 FIFA World Cup. *Tourism Management*, vol. 26, no. 4, pp. 595–603, 2005. doi:10.1016/j.tourman.2004.03.002.

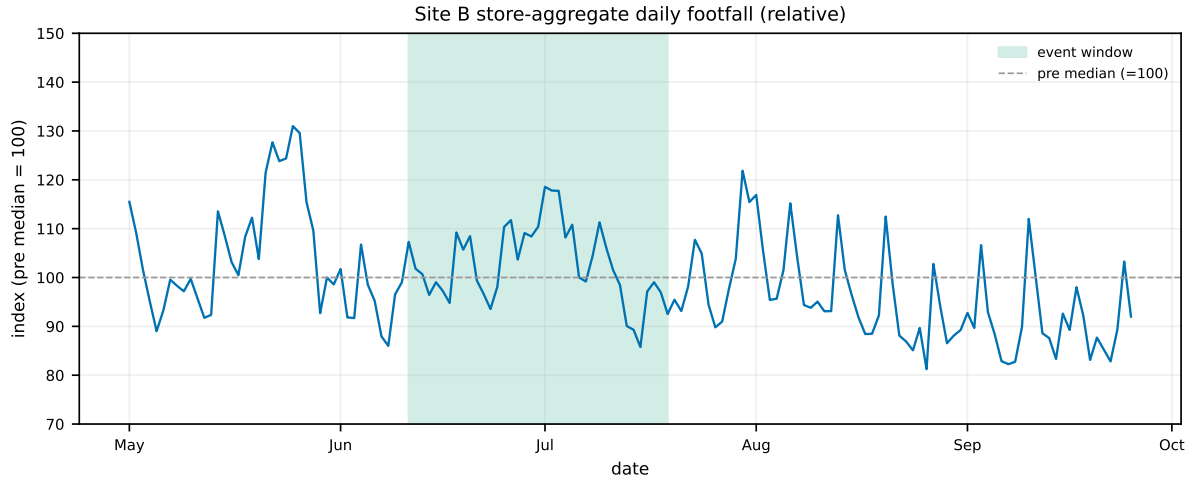


Figure 1: Site B store-aggregate daily footfall, indexed to the pre-window daily median (= 100), over the study window. The event window (shaded green) contains the mid-window hump; the post window reverts below the pre median. Volumes are relative (indexed) for anonymity; no raw magnitudes are shown. Numbers injected from `results/output.json` (Table 1, Section 4.1).

| estimate                      | value  | 95% CI lower | 95% CI upper | <i>p</i>  |
|-------------------------------|--------|--------------|--------------|-----------|
| store WC (vis/day)            | -2170  | -7895        | 3556         | 0.458     |
| store post (vis/day)          | -12019 | -17369       | -6670        | 1.065e-05 |
| zone-pooled WC (vis/zone-day) | -128   | -457         | 201          | 0.447     |
| placebo draws                 | 1486   | —            | —            | —         |
| placebo null mean (vis/day)   | -6372  | —            | —            | —         |
| placebo null s.d.             | 5297   | —            | —            | —         |
| placebo <i>P</i> (2-sided)    | 0.843  | —            | —            | —         |

Table 2: Primary day-of-week-adjusted contrast estimates (visitors/day), 95% CI, and *p* cluster-robust by calendar day, at the store aggregate and pooled across zones; placebo null and MDE. For readability we report MDE separately in Table 5. Generated from `results/output.json`.

| window       | days | median (raw) | median $\Delta$ vs pre (%) |
|--------------|------|--------------|----------------------------|
| event early  | 13   | 132222       | -0.6                       |
| event middle | 13   | 146831       | +10.3                      |
| event late   | 13   | 131033       | -1.5                       |
| post         | 68   | 123757       | -7.0                       |

Table 3: Event-window sub-period and post reversion contrasts. Generated from `results/output.json`.

| pre-window spec       | median $\Delta$ (%) | DOW-adj (vis/day) | 95% CI       | <i>p</i> | <i>n</i> |
|-----------------------|---------------------|-------------------|--------------|----------|----------|
| widened 41d (primary) | +1.5                | -2170             | [-7895,3556] | 0.458    | 148      |
| holiday-excluded      | +3.2                | 5013              | [854,9171]   | 0.018    | 137      |
| 10-day probe          | +5.9                | —                 | —            | —        | —        |

Table 4: Pre-baseline-robustness battery (DOW-adjusted and raw-median reads). Generated from `results/output.json`.

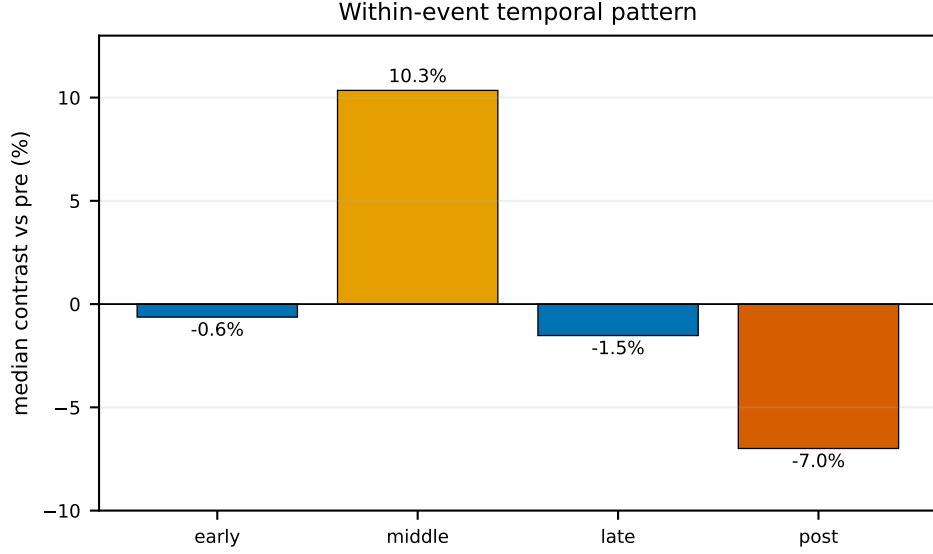


Figure 2: Event-window sub-periods (early / middle / late thirds) as raw-median contrast versus the pre-window median (%), with the post-window reversion. The middle third carries the clear positive hump; both neighbouring thirds are at or below pre levels, and the post window reverts below. Volumes are relative. Numbers injected from `results/output.json` (Table 3, Section 4.3).

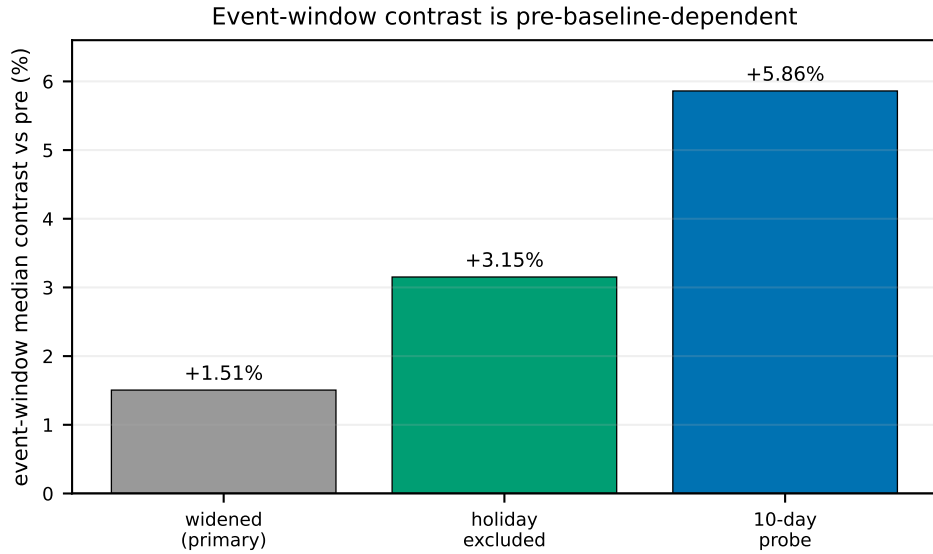


Figure 3: Pre-baseline-robustness battery: event-window raw-median contrast versus pre (%) under three pre-window specifications — widened 41-day (*widened, primary*), widened with the annual holiday peak excluded (*holiday-excluded*), and the original 10-day pre-window (*10-day probe*). The contrast is materially specification-dependent (see Table 4), the key honesty result. Values are relative. Numbers injected from `results/output.json`.

| MDE           | absolute (vis/day) | % of pre mean |
|---------------|--------------------|---------------|
| $\alpha=0.05$ | 5725               | 4.14          |
| 80% power     | 8184               | 5.92          |

Table 5: Minimum detectable effect at  $\alpha = 0.05$  and at 80% power, in absolute visitors/day and as % of the pre-window mean. Generated from `results/output.json`.