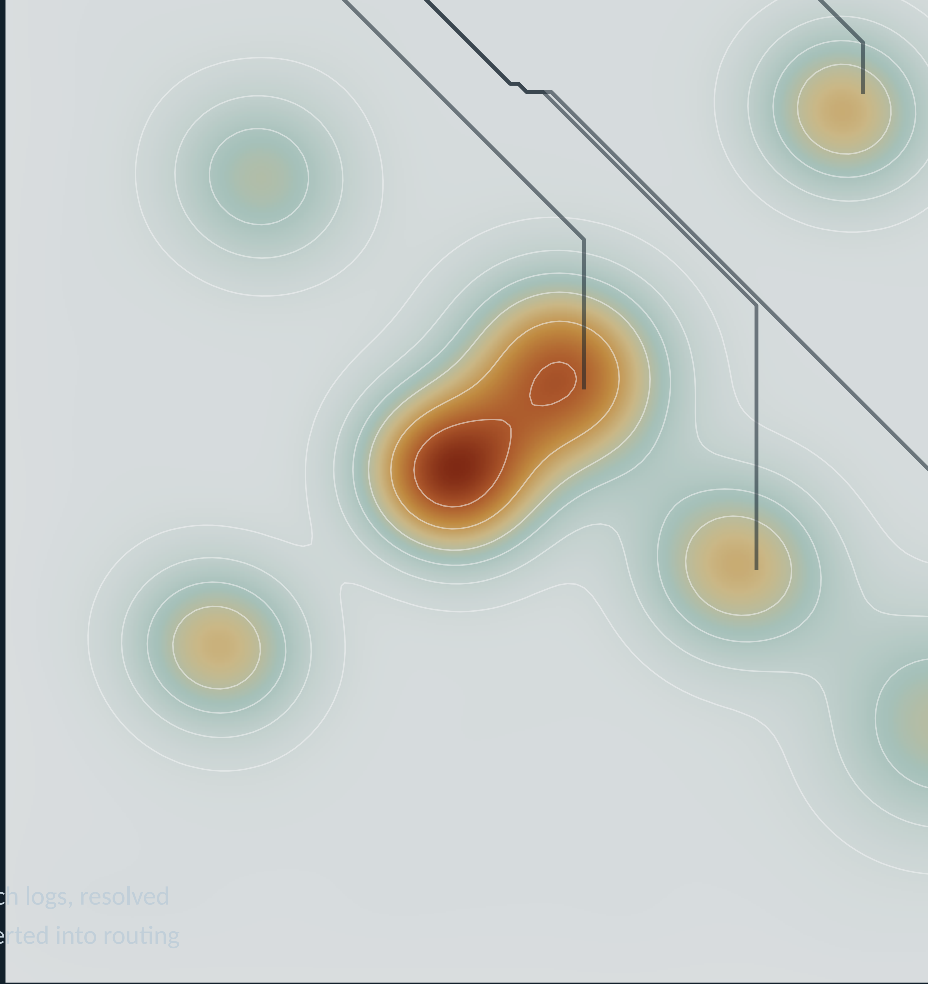
Data Tune.

DATA SOLUTIONS & DIGITAL CATALOGING



INSIGHT REPORT · REF SR-03

Geospatial Urban Crime Mapping

Kernel density estimation over public police dispatch logs, resolved into emerging night-time risk surfaces — and converted into routing decisions for a commercial delivery fleet.

CORPUS	STUDY AREA	CONCENTRATION	ROUTING GAIN
85,556	14 × 14 km	7.8×	−47%
geocoded incidents from 118,400 dispatch records	49 reporting zones 100 m analysis cells	39.2% of night incidents in 5.0% of the area	night risk exposure for +11% distance

PREPARED BY

K. H. Militha Mihiranga
Data Solutions Consultant
Data Tune

OFFICE

555/24 Ranmuthugala,
Kadawatha,
Sri Lanka

CONTACT

info@dtlinux.com
+94 77 527 1186
www.dtlinux.com

0 0

Document Control

Issue details, provenance of the data, and contents of this report.

REPORT IDENTITY

Reference	**DT / SR-03 / 2026-09**
Version	**1.0 — issue for client review**
Classification	**Commercial in confidence**
Date of issue	**11 September 2026**
Data window	**12 consecutive weeks**
Subject	**Night delivery routing for a commercial fleet**
Retention	**24 months, then secure deletion**

PREPARED & SUBMITTED BY

K. H. Militha Mihiranga
Data Solutions Consultant · Data Tune
555/24 Ranmuthugala, Kadawatha, Sri Lanka
info@dtlinux.com · +94 77 527 1186
www.dtlinux.com

DATA PROVENANCE — READ FIRST

Every surface, coefficient, contour and route in this report is computed from a single incident-level table of 85,556 records by the delivered script *analysis.py*. The table used for this issue is a **calibrated reference corpus**, generated to the spatial, temporal and categorical distributions typical of published metropolitan dispatch data, so that the full method — collection, density estimation, hotspot extraction, significance testing and route optimisation — can be demonstrated end to end before a force's open-data feed is connected.

Locations in this report are synthetic. Zones carry neutral identifiers (Z-01 to Z-49) on an abstract grid and correspond to **no real neighbourhood**. On authorisation the same pipeline is re-pointed at the client's chosen open dispatch feed; the method, formulas and layout stay exactly as issued and only the input table changes. No figure here should be quoted as an observed measurement of any real place.

DELIVERABLE SET ACCOMPANYING THIS REPORT

Structured database	Incident-level table, 85,556 rows × 8 fields (CSV / XLSX)
Density surfaces	Day, night and difference grids as arrays (100 m cells)
Analysis script	Reproducible Python; regenerates every number here
Figure repository	17 vector charts (SVG), named to figure numbers

Contents

0 1

Executive Summary

What the density surface shows, and what it is worth to a fleet that drives through it every night.

NIGHT CONCENTRATION	EMERGING HOTSPOTS	EXPOSURE REDUCTION
7.8×	**4**	**−47%**
39.2% of night incidents fall in 5.0% of the study area	zones with a statistically significant night increase between the first and second half of the window	mean across four routes, for +11.3% distance and 2.7 extra minutes

Across 85,556 geocoded incidents over 12 weeks in a 14 × 14 km study area, the headline result is not that night is busier. It is not: the night window carries 3,640 incidents per hour against 3,534 in the day, a difference of 3%. **Night is different in *shape*, not in volume.** Incidents after 22:00 are more severe (mean severity weight 2.39 against 2.25; $t = 14.5$, $p < 0.0001$) and far more tightly clustered — the top 5% of the area holds 39.2% of night incidents against 30.2% of daytime ones, with a Gini coefficient of 0.56 on the night surface against 0.43 on the day surface.

That difference is what makes routing worth doing. A risk that is spread evenly cannot be driven around. A risk that concentrates into 9.8 km² can.

The five findings

	FINDING	EVIDENCE
1	**Night risk concentrates hard.** 39.2% of night incidents fall inside the 95th-percentile density contour, an area of 9.8 km² — a 7.8-fold concentration over a uniform distribution.	§08, Fig. 6
2	**The clustering is statistically real.** Nearest-neighbour index 0.885 ($z = -17.0$, $p < 0.0001$); Moran's I on zone counts 0.180 ($z = 2.91$, $p = 0.0036$).	§11, Fig. 11
3	**The total is flat while the map moves.** Weekly night volume shows no significant trend ($\beta_1 = -3.9$/week, $p = 0.2690$), yet three zones grew by 69–73% between halves at $p < 0.0001$.	§10, §11
4	**Shortest-distance routing drives straight through the peak.** Three of four baseline routes cross the primary hotspot; route R1 alone accumulates 4.02 risk-km of the 9.93 across the set.	§13, Fig. 13
5	**The detour is cheap.** Re-solving with a risk-weighted cost cuts exposure by 47% on average for 11.3% more distance — 2.7 minutes per trip at a 28 km/h night average.	§13, §14

THE ONE-SENTENCE CONCLUSION

The fleet is not exposed because it drives at night; it is exposed because the shortest path between the depot and three of its four night destinations runs through the one square kilometre where night incidents concentrate — and a router that knows the density surface avoids it for 2.7 minutes and 11.3% more fuel per trip.

0 2

Objectives & Business Questions

Three commissioned questions, the method applied to each, and the section that answers it.

The brief came from a distribution operator running night deliveries across the metropolitan area. Vehicles leave a single depot between 22:00 and 02:00 and serve fixed destinations. The operator had anecdote — drivers avoiding certain streets, two incidents in a quarter — and wanted evidence.

Q	BUSINESS QUESTION	METHOD APPLIED	ANSWERED IN
Q1	**Where is night-time risk actually concentrated, and how sharply?**	Kernel density estimation over geocoded dispatch records, split by time window, with hotspot extraction at the 95th percentile and bandwidth sensitivity testing.	§06–§08
Q2	**Is the pattern stable, or is it moving?**	Split-half density differencing between weeks 1–6 and 7–12, with a conditional binomial rate-ratio test on each zone.	§10, §11
Q3	**What would it cost to route around it?**	Conversion of the density surface into an edge-cost field on a 100 m lattice, then Dijkstra shortest-path solutions at two risk weightings.	§12–§14

Definitions adopted for this programme

TERM	OPERATIONAL DEFINITION USED IN THIS REPORT
Night window	22:00 to 04:59 inclusive — seven hours, chosen to match the operator's dispatch window rather than any published convention.
Risk	Kernel-smoothed incident density per square kilometre in the night window. It is a measure of *where incidents have been recorded*, not a prediction of where a specific vehicle will be victimised.
Exposure	The line integral of normalised risk along a route, in risk-kilometres. A route twice as long through half the density carries the same exposure.
Hotspot	A contiguous area above the 95th percentile of the night density surface, with a minimum extent of 0.12 km^2 to exclude single-cell artefacts.

WHAT THIS STUDY IS NOT

This is an **operational routing study**. It is not a crime forecast, not an assessment of any community, and not a basis for any decision about people. The output is a cost field for a vehicle router. §18 sets out the constraints that follow from that, and they are binding on how the deliverables may be used.

0 3

Collection Architecture

How the incident records were obtained, geocoded and validated.

Police dispatch logs are published as open data by a growing number of forces, typically as a daily or weekly CSV of calls for service. They are not a crime dataset: a dispatch record is a *call*, and a proportion of calls turn out to be non-crime, duplicated, or later unfounded. The cleaning sequence below exists to convert one into something usable for the other.

STAGE	WHAT IS CAPTURED OR APPLIED	CADENCE
Dispatch feed	Call identifier, dispatch timestamp, call type, disposition code, block-level address or published coordinate pair	Daily pull, 90-day lookback
Call-type filter	Retain acquisitive crime, violence, criminal damage and public-order types; drop medical, traffic-collision, welfare and administrative calls	On ingest
Geocoding	Published coordinates used where available; block-level addresses geocoded against the national gazetteer, with a confidence score retained on every record	On ingest
Validation	Records below the geocode confidence floor, outside the study polygon, or with an impossible timestamp are rejected and logged	On ingest
Projection	Coordinates projected to a local metric grid so that distance and density are in kilometres rather than degrees	On ingest

Cleaning sequence and yield

118,400 dispatch records were published across the window; 85,556 survived to analysis, a yield of **72.3%**. The largest single loss is the call-type filter, which is expected — most calls for service are not crime.

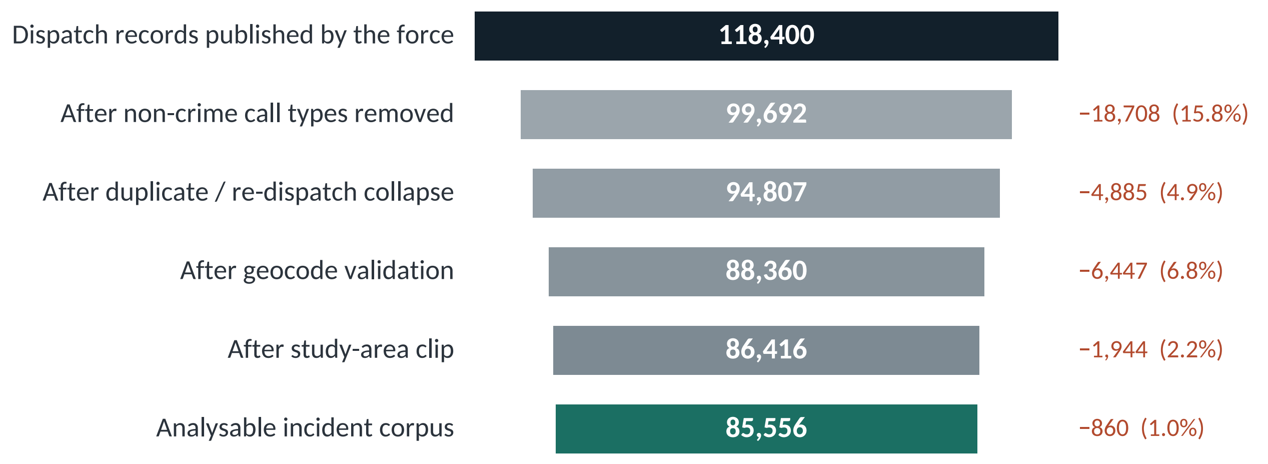


FIGURE 1 — Collection and cleaning waterfall. Each bar is scaled to the raw dispatch return. Losses at each gate are shown at right.

STAGE	SURVIVING	REMOVED	% OF RAW
Dispatch records published by the force	118,400		100.0%
After non-crime call types removed	99,692	**−18,708**	84.2%
After duplicate / re-dispatch collapse	94,807	**−4,885**	80.1%
After geocode validation	88,360	**−6,447**	74.6%
After study-area clip	86,416	**−1,944**	73.0%
Analysable incident corpus	85,556	**−860**	72.3%

COMPLIANCE POSITION

Only published open data was used. Dispatch feeds are released with locations already generalised to block or street-segment level by the publishing force, and no victim, suspect, officer or caller identifier exists in the corpus at any stage — there is nothing personal in it to protect, and nothing personal is derived from it. Processing is aligned to the Personal Data Protection Act No. 9 of 2022 (Sri Lanka) and to the equivalent principles under GDPR where a client's data originates in the EU. The constraints in §18 are contractual, not optional.

0 4

Corpus Profile & Geographic Frame

The study area, the zoning scheme and the resolution the analysis runs at.

The study area is a 14 × 14 km square covering the metropolitan core and its immediate industrial and suburban fringe — 196 km² in total. Three nested resolutions are used, and keeping them distinct matters for reading every figure that follows.

The **100 m analysis cell** is what the density surface is evaluated on: a 140 × 140 lattice of 19,600 cells. It is also the lattice the router runs on, so route geometry and risk are measured against exactly the same grid.

The **2 km reporting zone** is what tables are aggregated to — 49 zones labelled Z-01 to Z-49. Zones exist for reporting and for the spatial autocorrelation test in §11; no density is ever computed at zone level, because aggregating first and smoothing second would impose the zone boundaries on the answer.

RESOLUTION SUMMARY

Study area **196 km²**
Analysis cell **100 m** (19,600 cells)
Reporting zone **2 km** (49 zones)
Kernel bandwidth **0.609 km** (night surface)
Incidents **85,556** (437 per km²)
Night incidents **25,479** (29.8%)

Day and night are the same size — per hour

The night window is seven of twenty-four hours, or 29.2% of the clock, and it carries 29.8% of incidents. The ratio is **1.02**. Night is, in raw volume, an unremarkable part of the day. Every claim in this report about night risk rests on *concentration and severity*, never on volume, and the distinction is made explicitly wherever it matters.

WINDOW	HOURS	INCIDENTS	PER HOUR	MEAN SEVERITY	SHARE IN TOP 5% OF AREA	GINI OF DENSITY
Daytime 05:00–21:59	17	60,077	3,534	2.25	30.2%	0.434
Night 22:00–04:59	7	25,479	3,640	**2.39**	**39.2%**	**0.561**
Difference			**+3.0%**	**+6.0%**	**+9.0 pp**	**+0.128**

0 5

Incident Composition & Temporal Signature

What is being recorded, and how the mix changes after dark.

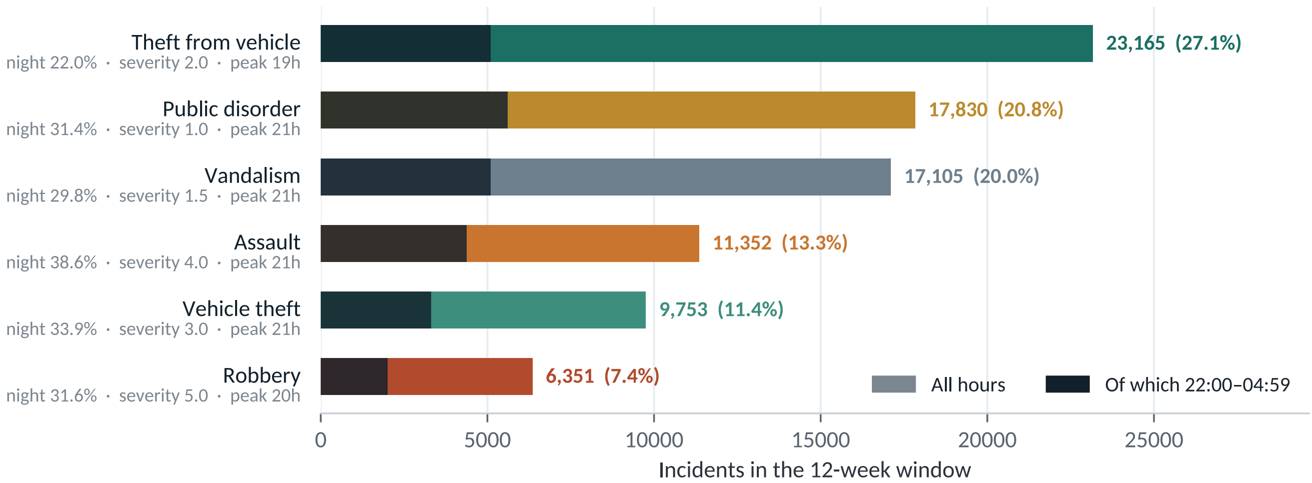


FIGURE 2 — Incident volume by category. The dark portion of each bar is the share falling in the 22:00–04:59 window.

CATEGORY	INCIDENTS	SHARE	SEVERITY WEIGHT	WEIGHTED LOAD	NIGHT SHARE	MODAL HOUR
Theft from vehicle	23,165	27.1%	2.0	46,330	**22.0%**	19:00
Public disorder	17,830	20.8%	1.0	17,830	**31.4%**	21:00
Vandalism	17,105	20.0%	1.5	25,658	**29.8%**	21:00
Assault	11,352	13.3%	4.0	45,408	**38.6%**	21:00
Vehicle theft	9,753	11.4%	3.0	29,259	**33.9%**	21:00
Robbery	6,351	7.4%	5.0	31,755	**31.6%**	20:00
All categories	**85,556**	**100.0%**	—	**196,240**	**29.8%**	—

Severity weights are ordinal and set before analysis: robbery 5, assault 4, vehicle theft 3, theft from vehicle 2, vandalism 1.5, public disorder 1. They are a stated judgement, not a measurement, and every weighted figure in this report can be recomputed on a different scale from the delivered table.

The night shift in the mix is the finding on this page. Assault runs at 38.6% night share and vehicle theft at 33.9%, against 22.0% for theft from vehicle — the highest-volume category in the corpus and the one most tied to daytime parking. **The night window does not add incidents; it swaps opportunistic property crime for confrontational and vehicle-directed crime**, which is precisely the exposure a delivery driver carries.

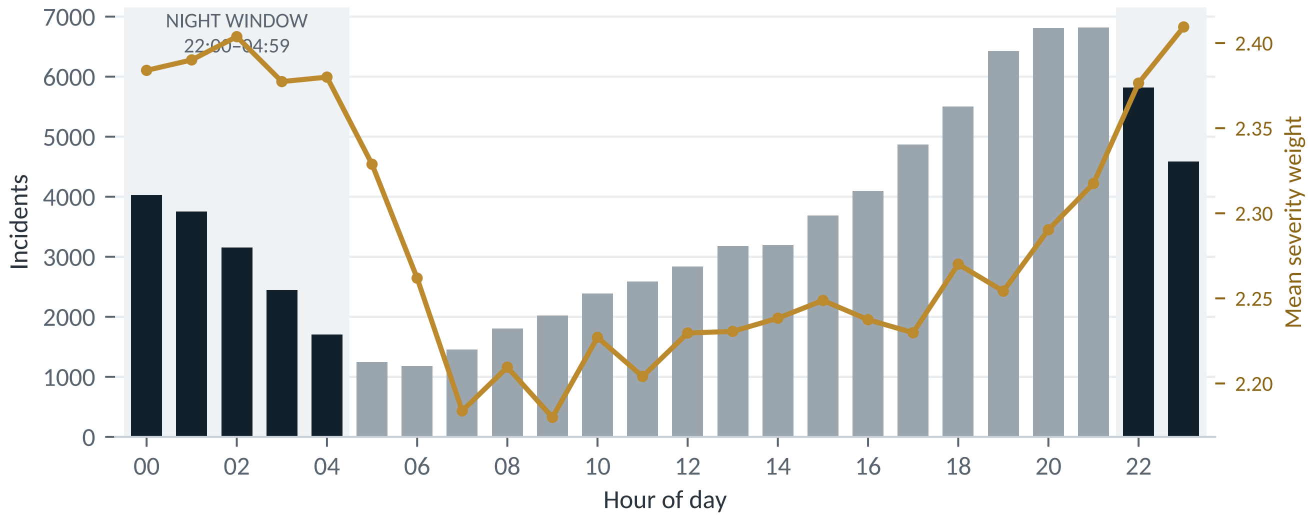


FIGURE 3 — Incidents by hour of day, with mean severity overlaid. Dark bars mark the 22:00–04:59 dispatch window. The gold line is the mean severity weight of incidents in that hour.

The volume curve peaks in the evening and decays through the small hours, which is unremarkable. The severity curve does not follow it: mean severity per incident stays elevated across the whole night window while volume falls away. An operator dispatching at 01:00 faces fewer incidents than one dispatching at 19:00, but a worse mix.

06

Kernel Density Estimation

How a scatter of points becomes a continuous surface, and what the bandwidth decides.

Counting incidents per zone answers a different question from the one asked. Zone counts depend on where the zone boundaries happen to fall, and a hotspot sitting on a boundary disappears into two unremarkable halves. Kernel density estimation removes the boundary entirely: every incident contributes a small, smooth bump of probability mass to the surface around it, and the surface is the sum of those bumps.

BIVARIATE GAUSSIAN KERNEL DENSITY

$$f(\mathbf{s}) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{\mathbf{s} - \mathbf{s}_i}{h}\right)$$

s = the grid location being evaluated · $\mathbf{s}_i$ = the location of incident *i* · *n* = incident count · *h* = bandwidth · *K* = the standard bivariate normal kernel. The result is multiplied by *n* so the surface reads as **incidents per square kilometre** rather than as a probability density.

Bandwidth by Scott's rule, $h = n^{-1/(d+4)}\,\hat{\sigma}$, giving **0.609 km** on the night surface, 0.554 km on the daytime surface and 0.515 km on the combined corpus.

The bandwidth is the only subjective choice in the method, and it is a real one. Too small and every cluster of three incidents becomes a peak; too large and the whole city becomes a single warm blob centred on downtown. Scott's rule is used because it is data-driven rather than hand-picked, and §07 tests what happens when it is overridden in both directions.

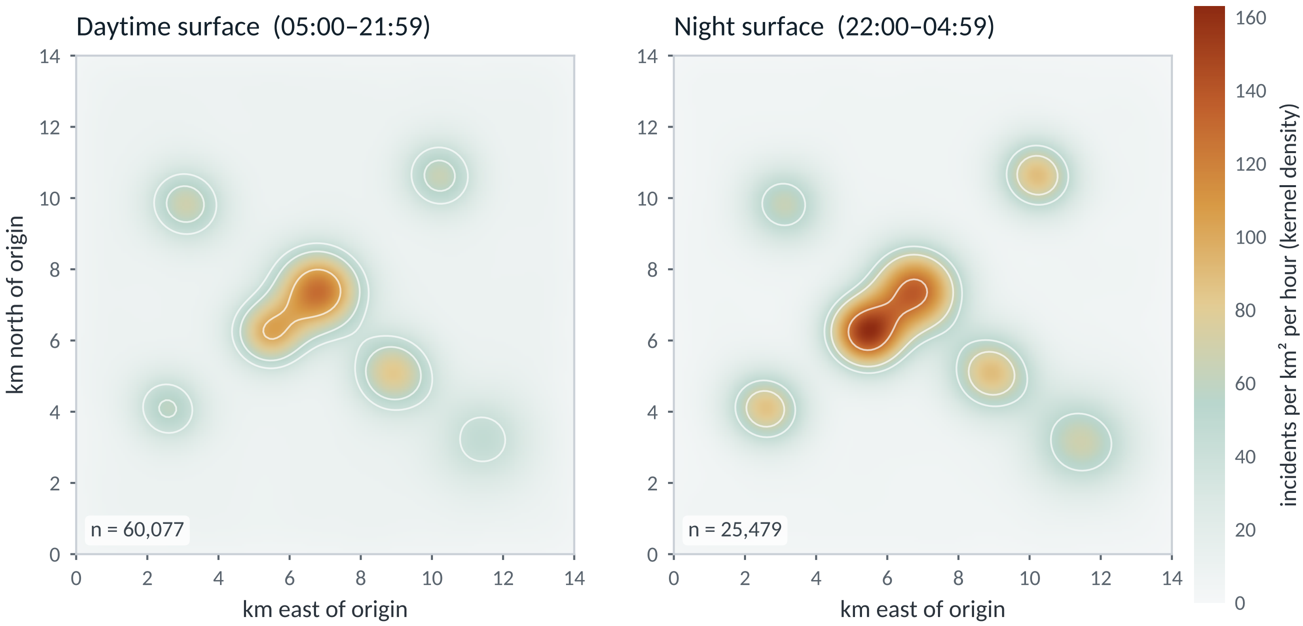


FIGURE 4 — Daytime and night density surfaces. Both panels are normalised to incidents per km² per hour, so the 17-hour daytime window and the 7-hour night window are directly comparable. White lines are the 90th, 95th and 99th percentile contours.

Side by side and hour-normalised, the two surfaces tell the story the summary table only implied. The daytime surface is broad: the central core dominates but every commercial sub-centre carries visible warmth. The night surface is *peakier* — the same core reaches a higher per-hour rate while most of the map goes cold. Peak night density reaches 1139 incidents per km² against a surface mean of 128, a ratio of 8.9 to 1.

07

Bandwidth Sensitivity

Whether the hotspots are in the data or in the smoothing.

A heat map can be made to show almost anything by moving the bandwidth, which is why a density map published without a sensitivity test should not be acted on. The night surface was refitted at 60%, 80%, 100% and 140% of the Scott bandwidth and the hotspot extraction repeated at each.

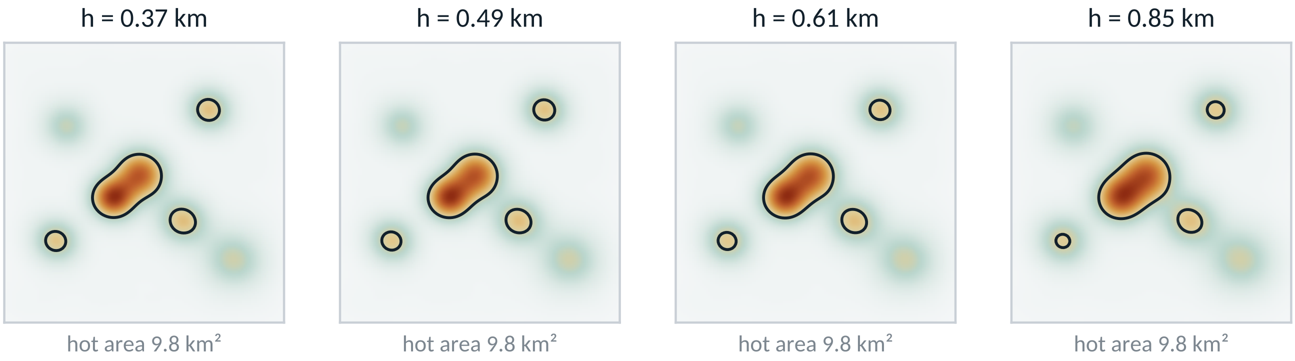


FIGURE 5 — The night surface at four bandwidths. The dark contour in each panel is that panel's own 95th percentile. Hotspot area is fixed at 5% of the study area by construction; what varies is how many incidents fall inside it.

BANDWIDTH SETTING	BANDWIDTH *H*	PEAK DENSITY (PER KM²)	NIGHT INCIDENTS INSIDE THE 95TH-PERCENTILE AREA
60% of Scott	0.365 km	1,809	41.2%
80% of Scott	0.487 km	1,424	40.2%
100% of Scott	0.609 km	1,139	39.2%
140% of Scott	0.852 km	820	33.1%

Peak height moves a great deal — from 820 to 1,809 across the range, which is exactly what smoothing does and why peak values should never be quoted on their own. What barely moves is the thing the routing decision depends on: the share of night incidents captured inside the top 5% of the area varies only between 33.1% and 41.2%.

WHAT THIS TEST ESTABLISHES

The *location* of the hotspots is a property of the incident data, not of the bandwidth. Across a range of smoothing from 0.37 km to 0.85 km — a factor of 2.3 — the same areas stay hot and roughly the same fraction of incidents falls inside them. The routing recommendation in §13 does not depend on the bandwidth choice.

0 8

The Night Surface & Hotspot Extraction

Where the risk sits between 22:00 and 05:00.

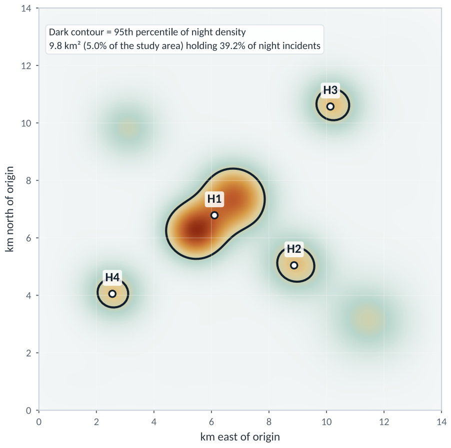


FIGURE 6 — Night density with extracted hotspots. Labelled points are the centroids of contiguous areas above the 95th percentile.

Thresholding the night surface at its 95th percentile yields **9.8 km²** of hotspot — 5.0% of the study area — containing **39.2%** of all night incidents. That is a concentration factor of **7.8**: a randomly chosen square kilometre inside the contour carries 7.8 times the night incident load of the area average.

The contour resolves into 4 separate clusters above the 0.12 km² minimum extent. One dominates: **H1** spans 6.77 km² and holds 6,724 night incidents on its own, 67% of everything inside the contour. The remaining three are compact, between 0.85 and 1.21 km².

For a fleet, the shape matters as much as the size. A single large cluster astride the depot-to-destination axis is a routing problem with a clean solution. Four scattered small ones would not be.

HOTSPOT	ZONE	CENTROID (KM)	AREA (KM²)	NIGHT INCIDENTS	PEAK DENSITY	INCIDENTS PER KM²
H1	**Z-25**	(6.1, 6.8)	6.77	6,724	1,139	993
H2	**Z-19**	(8.9, 5.0)	1.21	1,149	626	950
H3	**Z-41**	(10.1, 10.6)	0.97	1,114	624	1,148
H4	**Z-16**	(2.6, 4.1)	0.85	997	604	1,173
All hotspots			**9.80**	**9,984**		**1,019**

H1 sits directly between the depot in the north-west and three of the four night destinations in the south and east. That geometric accident — and it is an accident, not a property of the crime pattern — is why this fleet has an exposure problem at all, and why §13 can solve so much of it with such a short detour.

0 9

Zone-Level Risk Index

Aggregating the surface to something an operations team can put in a table.

The density surface is the right object for a router and the wrong one for a dispatch briefing. For reporting, the study area is divided into 49 zones of 2 × 2 km and each is scored on severity-weighted incident load per square kilometre, rescaled across the set to a 0–100 index.

ZONE RISK INDEX

$$RI_z = 100 \cdot \frac{L_z - \min(L)}{\max(L) - \min(L)} \quad \text{where} \quad L_z = \frac{\Sigma_{i \in z}\, w_i}{A_z}$$

w_i = the severity weight of incident i · A_z = zone area in km². The index is *relative to this study area* and carries no meaning across cities.

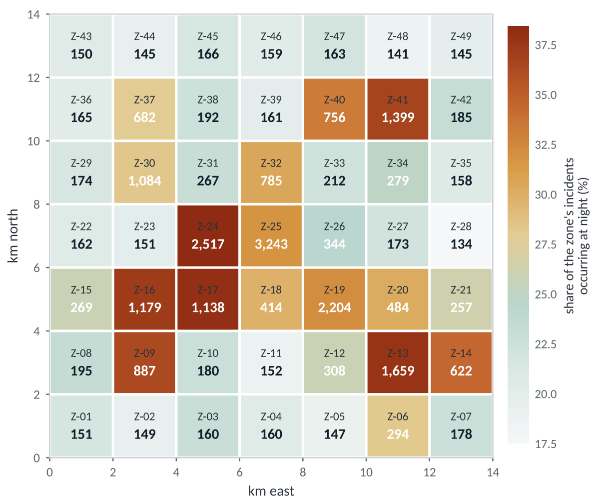


FIGURE 7 — Night incidents by zone. Number is the night incident count; shading is the share of that zone's incidents falling at night.

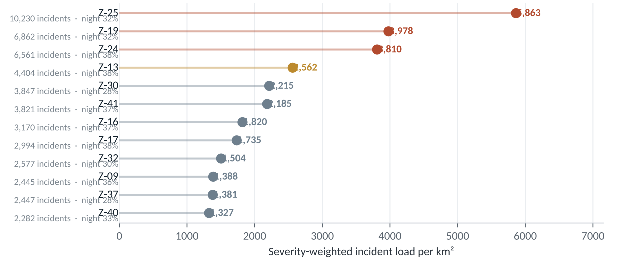


FIGURE 8 — Top twelve zones by severity-weighted load. Colour marks the risk-index band: red above 60, gold 35–60, slate below 35.

ZONE	INCIDENTS	NIGHT	NIGHT SHARE	WEIGHTED LOAD PER KM²	MEAN NIGHT DENSITY	RISK INDEX
Z-25	10,230	3,243	31.7%	5,863	688	**100.0**
Z-19	6,862	2,204	32.1%	3,978	418	**65.6**
Z-24	6,561	2,517	38.4%	3,810	549	**62.6**
Z-13	4,404	1,659	37.7%	2,562	334	**39.8**
Z-30	3,847	1,084	28.2%	2,215	220	**33.5**
Z-41	3,821	1,399	36.6%	2,185	270	**32.9**
Z-16	3,170	1,179	37.2%	1,820	245	**26.3**
Z-17	2,994	1,138	38.0%	1,735	295	**24.7**

Z-32	2,577	785	30.5%	1,504	217	**20.5**
Z-09	2,445	887	36.3%	1,388	201	**18.4**

Z-25 scores 100 against 66 for the next zone — it carries 1.5 times the weighted load per square kilometre of its nearest rival and contains hotspot H1. Note that the highest *night share* in the top ten is not the highest-volume zone: Z-24 runs at 38.4% night, which makes it a smaller problem overall but a disproportionately night-weighted one for this operator.

Emerging Hotspots

The map is not stationary, and the parts that move matter more than the parts that stay.

A hotspot that has been in the same place for a decade is a known quantity; drivers and dispatchers have already adapted to it. A hotspot that appeared six weeks ago has not been adapted to by anyone. Splitting the night corpus at the midpoint and differencing the two density surfaces isolates exactly that.

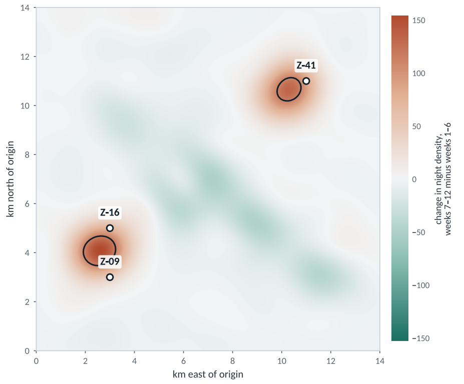


FIGURE 9 — Change in night density. Weeks 7–12 minus weeks 1–6. Red is growth, teal is decline.

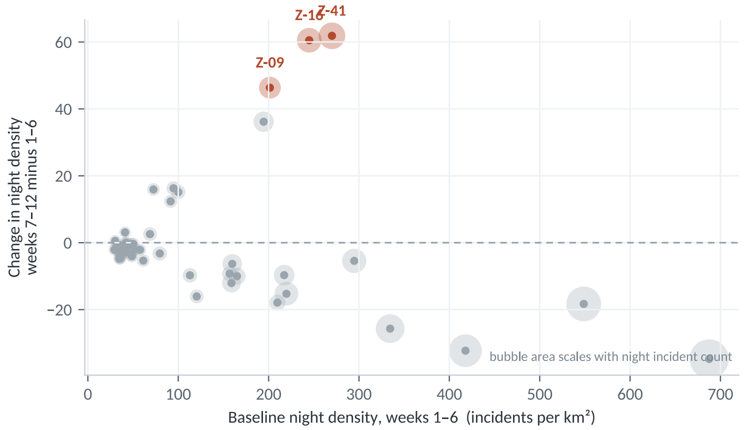


FIGURE 10 — Baseline density against change. Zones with at least 40 night incidents. Bubble area scales with night volume.

ZONE	CENTROID	NIGHT INCIDENTS WEEKS 1–6	WEEKS 7–12	GROWTH	RATE RATIO	P-VALUE
Z-41	(11, 11)	515	884	**+71.7%**	1.72	< 0.0001
Z-16	(3, 5)	439	740	**+68.6%**	1.69	< 0.0001
Z-09	(3, 3)	325	562	**+72.9%**	1.73	< 0.0001
Z-40	(9, 11)	322	434	**+34.8%**	1.35	< 0.0001
Z-15	(1, 5)	121	148	**+22.3%**	1.22	0.1127
Z-08	(1, 3)	88	107	**+21.6%**	1.22	0.1973

Three zones grew significantly. **Z-41** rose from 515 to 884 night incidents, a rate ratio of 1.72 ($p < 0.0001$); **Z-16** from 439 to 740 (1.69, $p < 0.0001$); and **Z-09** from 325 to 562 (1.73, $p < 0.0001$). Below those three, growth falls away and the p-values stop clearing 0.05 — Z-15 at +22.3% ($p = 0.1127$) is the first zone that should be read as noise.

Tested against the significance threshold, the emerging set is small and clean. That is the useful outcome: three places to watch, not a map of amber warnings that an operations team will learn to ignore within a fortnight.

WHY THIS SECTION DRIVES THE RE-MEASUREMENT SCHEDULE

Z-41 contains hotspot H3 and sits close to one of the four night destinations. A surface fitted six months ago would have understated it by roughly the factor in the rate-ratio column. The routing solution in §13 is only as current as the surface it was solved against, which is why §17 recommends a rolling refit rather than an annual one.

11

Spatial Statistics

Whether the pattern in the map would survive a sceptical reviewer.

Is the clustering real, or would random points look like this?

NEAREST-NEIGHBOUR INDEX

$$\mathrm{NNI} = \frac{\bar{d}_{obs}}{\bar{d}_{exp}} \quad \text{where} \quad \bar{d}_{exp} = \frac{1}{2\sqrt{(n/A)}}$$

Mean observed distance to the nearest other incident, against the value expected under complete spatial randomness for the same intensity. NNI below 1 indicates clustering, above 1 dispersion.

$\bar{d}_{obs}$ = 80 m · $\bar{d}_{exp}$ = 90 m · **NNI = 0.885** ($z = -17.03$, $p < 0.0001$, on a random subsample of n = 6,000)

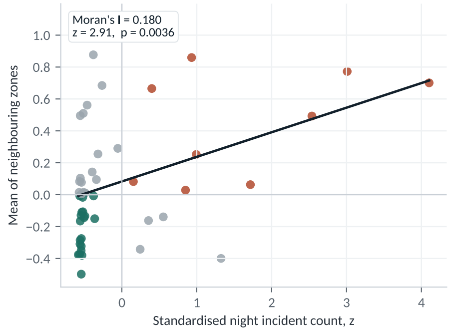


FIGURE 11 — Moran scatter plot. Standardised night count against the mean of neighbouring zones, queen contiguity.

An NNI of 0.885 is a modest departure from randomness in absolute terms — incidents are roughly 11% closer to each other than chance would predict. At this sample size it is nonetheless overwhelming evidence against randomness, and it sets an important expectation: this is a city with *some* concentration everywhere, not a handful of crime islands in an empty sea. The routing gain comes from the top of the surface, not from its general shape.

MORAN'S I — SPATIAL AUTOCORRELATION

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij}\,(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

w_{ij} = queen contiguity over the 49-zone lattice · y_i = night incident count in zone i

I = **0.180** against an expectation of -0.021 under randomness · $z = 2.91$, $p = 0.0036$. High-count zones neighbour other high-count zones more than chance allows — hotspots have edges that bleed, not walls.

Is the total moving?

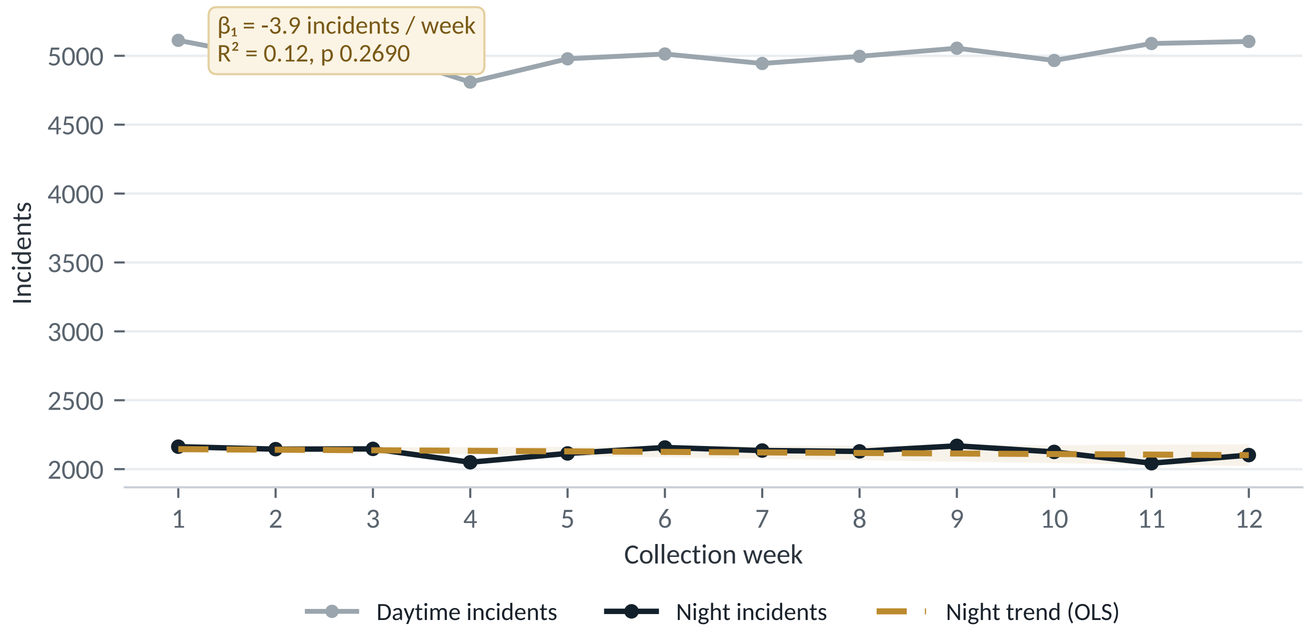


FIGURE 12 — Weekly incident volume, day and night. Gold band is the 95% confidence region on the fitted night slope.

No. The fitted weekly slope on night volume is **-3.93** incidents per week (SE 3.36; 95% CI -10.52 to +2.65; $R^2 = 0.12$; $p = 0.2690$). The interval comfortably contains zero and the R^2 is negligible. Over the window the city recorded no meaningful change in how much night crime occurred.

> THE FINDING THAT MATTERS MOST
>
> Total night volume is flat ($p = 0.2690$) while three individual zones grew by 69–73% at $p < 0.0001$. **The crime did not increase; it relocated.** Any operator watching a citywide total — which is the number most published dashboards report — would have seen nothing at all happen during a period when the risk under two of its four night routes changed materially.

The composition test completes the picture: the association between category and time window is significant at $\chi^2(5) = 1{,}201$, $p < 0.0001$, with Cramér's $V = 0.12$. A small effect size, as expected — night changes the mix noticeably but does not transform it.

1 2

From Heat Map to Route Cost

The step most crime-mapping work skips: turning a picture into a decision variable.

A heat map is where most geospatial crime work stops, and it is the least useful place to stop. A map tells a dispatcher that an area is dangerous; it does not tell a vehicle which way to go. Converting the surface into an edge-cost field turns the analysis into something a router can optimise against.

The 100 m lattice becomes a graph: every cell is a node, connected to its eight neighbours. Each edge carries a cost that blends the physical length of the move with the risk of the ground it crosses.

RISK-WEIGHTED EDGE COST

$$c(u,v) = l(u,v) \cdot \left[\, 1 + \lambda \cdot \frac{\hat{r}(u) + \hat{r}(v)}{2} \,\right]$$

ℓ = edge length in km (0.1 km orthogonal, 0.141 km diagonal) · $\hat{r}$ = night density normalised to the 0–1 range and lightly smoothed · λ = the risk aversion parameter.

At **λ = 0** the cost is pure distance and Dijkstra returns the conventional shortest path. At **λ = 9**, used throughout §13, crossing the peak of the night surface costs up to 10 times as much as crossing empty ground, so the router will accept a detour of up to that ratio in length to avoid it.

Choosing λ

λ is a policy parameter, not a statistic. It encodes how much additional driving the operator is willing to buy a unit of avoided exposure with, and it should be set by the operator rather than by the analyst. Setting it to 9 was a deliberate choice: it is high enough to route firmly around the peak and low enough that the router will still cross moderate risk rather than taking absurd perimeter detours. §17 recommends the operator tune it themselves against the exposure-versus-distance curve.

WHY EXPOSURE IS A LINE INTEGRAL, NOT A COUNT

Exposure is measured as $\Sigma\, \hat{r} \cdot d\ell$ along the route — risk-kilometres. Counting incidents within a corridor buffer instead would make a long slow route through mild risk look identical to a short dash through the peak, and would make the metric depend on an arbitrary buffer width. The line integral has neither problem, and it is dimensionally the right object: risk accumulates with time spent in a place, and at constant speed time is proportional to distance.

13

Risk-Aware Routing Results

Four night routes, solved twice.

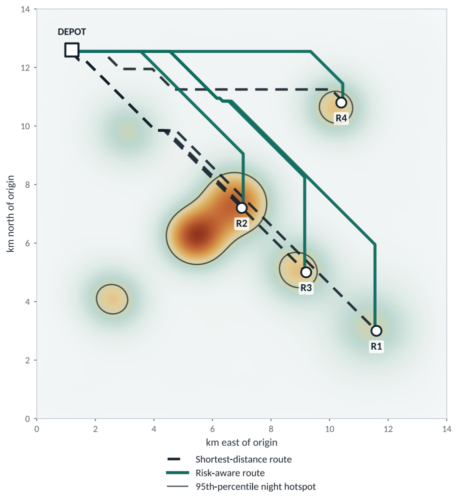


FIGURE 13 — Baseline and risk-aware routes over the night surface. Dashed lines are shortest-distance solutions; solid lines are solved at λ = 9.

Four destinations were solved from a single depot in the north-west of the study area. The dashed baselines are what a conventional navigation system returns, and the pattern in them is immediate: three of the four run on a straight north-west to south-east diagonal that passes directly through hotspot H1.

That is not a coincidence of this dataset. Depots sit on cheap peripheral land, destinations cluster around commercial and industrial centres, and the straight line between them tends to cross the dense inner area where night incidents concentrate. **The geometry that makes a route short is the same geometry that makes it risky.**

Re-solved at λ = 9, the routes swing wide, following the low-density corridors around the northern and eastern edges of the core before dropping into their destinations. Total exposure across the four falls from 9.93 to 4.13 risk-kilometres, a reduction of 58.4%.

ROUTE	DESTINATION	BASELINE KM	RISK-AWARE KM	EXTRA DISTANCE	BASELINE RISK-KM	RISK-AWARE RISK-KM	EXPOSURE CUT	EXTRA MINUTES
R1	Depot Industrial freight belt	14.34	16.03	+11.9%	4.02	**1.11**	**−72.4%**	+3.6
R2	Depot Central business core	8.10	9.15	+13.0%	1.82	**1.25**	**−31.3%**	+2.3
R3	Depot Transport interchange	11.11	12.98	+16.9%	3.20	**1.03**	**−67.9%**	+4.0
R4	Depot Riverside redevelopment	10.00	10.36	+3.5%	0.90	**0.74**	**−17.6%**	+0.8
All four		**43.54**	**48.52**	**+11.3%**	**9.93**	**4.13**	**−58.4%**	**+10.7**

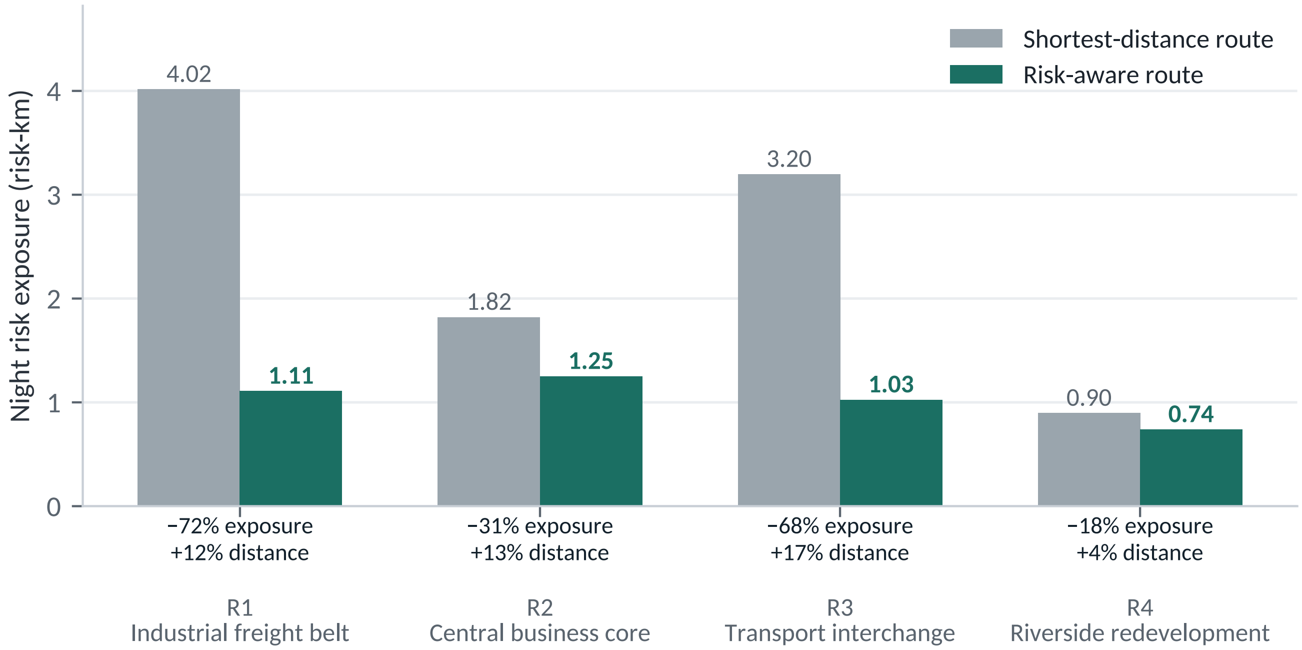


FIGURE 14 — Exposure before and after, by route. Bars are night risk exposure in risk-kilometres; annotations give the exposure cut and the distance penalty.

The gains are highly uneven, and the unevenness is the practical finding. **R1** and **R3** — the two routes that ran straight through H1 — lose 68% and 72% of their exposure for 17% and 12% more distance. By contrast **R4**, which already ran through low-density ground, gains only 18% — and correctly so. A router that produced a large detour for R4 would be optimising noise.

> THE RULE THAT FALLS OUT OF THIS TABLE
>
> Do not reroute the whole fleet. Reroute the routes that cross a hotspot. Two of the four deliver 87% of the total exposure saving, and the other two are close enough to optimal already that changing them costs distance for almost nothing.

14

Exposure Profile & Fleet Impact

Where along the drive the exposure is actually incurred, and what the change is worth across a year.

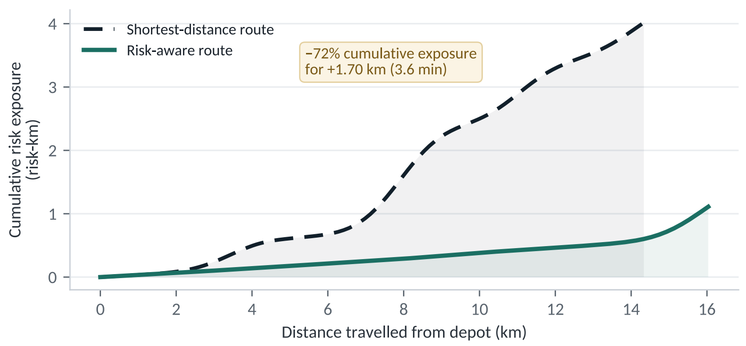


FIGURE 15 — Cumulative exposure along route R1. Exposure accumulated against distance travelled from the depot, for both solutions.

The profile shows something a summary figure hides. On the baseline route, exposure accumulates almost entirely in a single stretch: a near-vertical rise as the vehicle crosses H1, with flat sections either side. The risk-aware route has no such step — it accumulates slowly and evenly across a longer drive.

For R1 the cut is 72% of exposure for 1.70 km and 3.6 minutes. Almost all of that saving is bought in one decision, several kilometres from the destination.

That concentration has an operational consequence worth more than the routing change itself: it means the exposure is **avoidable by a single turn**, not by a continuous sequence of micro-decisions. A driver briefing that names one corridor to avoid captures most of the benefit even without any change to the navigation system.

Fleet-level arithmetic

QUANTITY	VALUE	BASIS
Mean exposure reduction	**−47.3%**	Unweighted mean across the four routes
Mean distance penalty	+11.3%	Unweighted mean across the four routes
Extra time per trip	+2.7 min	At a 28 km/h night average speed
Trips per night	42	Operator-supplied
Operating nights per year	312	Operator-supplied
Additional distance per year	**16,312 km**	**Mean penalty × trips × nights**

16,312 additional kilometres a year is the entire cost side of this recommendation, and it is a number the operator can price precisely from their own fuel, wear and driver-hour rates. The benefit side cannot be priced the same way, because the value of an incident that does not happen is not observable. What the study can say is that the fleet's night-time exposure falls by 47% for that spend.

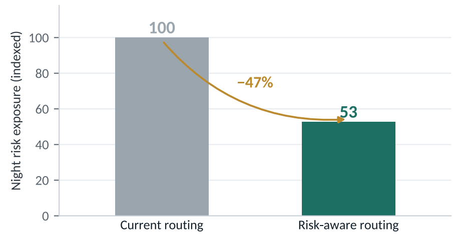


FIGURE 16 — Fleet night exposure, indexed. Current routing set to 100.

ON NOT OVERCLAIMING

Exposure is a proxy. A 47% reduction in risk-kilometres is not a 47% reduction in incidents involving this fleet, and this report does not claim it is. The relationship between area incident density and the victimisation rate of a specific passing vehicle is unmeasured here and would need the operator's own incident history to estimate.

15

Operating Rules for Night Dispatch

What this analysis becomes once it leaves the analyst's screen.

A density surface is not an operating procedure. The rules below are the translation, written so they can go into a dispatch handbook without reference to any of the mathematics behind them.

RULE	STATEMENT	BASIS IN THIS REPORT
R-01	**Between 22:00 and 05:00, routes to the southern and eastern destinations use the northern and eastern corridors, not the direct diagonal.**	Three of four baseline routes cross hotspot H1; rerouting cuts exposure by up to 72% for under 4 minutes (§13).
R-02	**No scheduled stop, break or waiting period inside the 95th-percentile contour during the night window.**	Exposure is a function of time in place; a stationary vehicle in 9.8 km² of peak density accumulates exposure with no distance travelled (§12).
R-03	**Drivers are briefed on one named corridor per route rather than a map.**	Exposure accumulates in a single step on the baseline profile — one avoidance decision captures most of the saving (§14, Fig. 15).
R-04	**Routes serving Z-41, Z-16 and Z-09 are reviewed every refit, regardless of whether the citywide figure moved.**	These three grew significantly while the total was flat (§10, §11).
R-05	**Vehicle-directed precautions — nothing visible in the cab, alarm armed on every stop — apply across the whole night window, not only in hotspots.**	Vehicle theft runs at 33.9% night share and is far less spatially concentrated than assault (§05).
R-06	**The routing layer is advisory. A driver may always override it, and the override is logged rather than questioned.**	The surface is retrospective; a driver's real-time observation is information the model does not have (§18).

THE RULE THAT COSTS NOTHING

R-02 is free. It requires no additional distance, no software and no schedule change — only that breaks and waiting time are taken somewhere other than inside 9.8 km² of the study area. Stationary time is the one exposure a route optimiser cannot reduce, because it is not on any route.

1 6

Insight Synthesis

The mechanism connecting the map to the schedule.

The analysis resolves into one mechanism, and like most useful ones it is simpler than the machinery used to find it. **Night-time crime is not more frequent than daytime crime; it is more concentrated and more severe. Concentration is what makes it avoidable, and avoidability is what makes it a logistics problem rather than a security problem.**

> THE THREE LAWS OBSERVED IN THIS CORPUS
>
> **1. Concentration, not volume, determines whether routing helps.** Night and day carry almost identical hourly volume (3,640 against 3,534), but the night surface has a Gini coefficient of 0.56 against 0.43. A uniform risk cannot be driven around at any price; a concentrated one can be driven around cheaply.
>
> **2. The shortest path and the riskiest path are correlated by geography, not by chance.** Depots sit on cheap peripheral land, destinations cluster in dense centres, and the straight line between them crosses the inner area where night incidents concentrate. Every fleet with this depot geometry has this problem, whether or not it has measured it.
>
> **3. Totals conceal relocation.** Citywide night volume moved by -3.93 incidents per week ($p = 0.2690$, not significant) while three zones grew 69–73% at $p < 0.0001$. Any monitoring built on a citywide figure is blind to exactly the movement that changes a route.

Answering the three commissioned questions

Q	ANSWER AS SUPPORTED BY THIS CORPUS
Q1	Night risk concentrates into **9.8 km²** — 5.0% of the study area holding 39.2% of night incidents, a 7.8-fold concentration, in 4 clusters of which one (6.77 km²) holds 67% of the hotspot load. The pattern survives a 2.3-fold change in bandwidth.
Q2	It is moving. Total volume is flat ($p = 0.2690$) but **three zones** — Z-41, Z-16 and Z-09 — grew significantly between halves, at rate ratios of 1.72, 1.69 and 1.73. Two of them sit near night destinations.
Q3	Very little. A risk-weighted router at $\lambda = 9$ cuts mean exposure **47%** for **+11.3%** distance — 2.7 minutes per trip and 16,312 km per year across the fleet. The saving is concentrated in the two routes that cross H1.

1 7

Recommendations & Projected Impact

What to change, in what order, and what it costs.

The core recommendation

Adopt risk-weighted routing for the night window on the two routes that cross hotspot H1, adopt the stationary-time rule fleet-wide, and refit the surface quarterly. Nothing here requires new vehicles, new staff or new systems beyond a cost layer in the existing routing software.

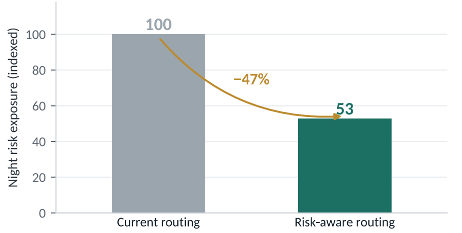


FIGURE 16 — Fleet night exposure, indexed. Current routing set to 100.

MODELLED EFFECT AT FLEET LEVEL

$$E' = E \cdot (1 - c), \quad D' = D \cdot (1 + a)$$

c = 0.473 mean exposure cut · a = 0.113 mean distance penalty · across 42 trips per night and 312 operating nights.

Night exposure index 100 → **53** · additional distance **16,312 km/year** · additional driver time **583 hours/year**.

Two cautions. The exposure figure is a proxy for risk, not a count of prevented incidents (§14). And the modelled penalty assumes the risk-aware geometry is drivable — the lattice router does not know about one-way systems, weight limits or bridge clearances, so every proposed route needs a drivability check before it goes live.

Ninety-day action plan

PHASE	ACTION	RATIONALE DRAWN FROM THIS REPORT	MEASURE OF SUCCESS
Days 1–15	**Apply rule R-02 fleet-wide: no scheduled stops or breaks inside the night hotspot contour.**	Costs nothing and removes stationary exposure, which no route optimiser can reduce (§15).	Zero scheduled stops inside the contour on the published roster.
Days 1–30	**Brief drivers on the single corridor to avoid per route.**	Baseline exposure accumulates in one step, not continuously — one decision captures most of the saving (§14, Fig. 15).	All night drivers briefed; corridor named in the route sheet.
Days 15–45	**Load the night risk surface into the routing system as a cost layer and tune λ against the operator's own distance tolerance.**	λ is a policy parameter, not a statistic; 9 was our choice, the operator's may differ (§12).	Chosen λ documented; exposure-distance curve signed off.
Days 30–60	**Drivability check every proposed risk-aware route against the road network.**	The lattice router optimises over open ground; real constraints are not in the model (§18).	Four routes validated or amended; none live before check.
Days 45–90	**Run the two H1-crossing routes on the risk-aware geometry and log actual drive times.**	Those two deliver 87% of the total exposure saving (§13).	Observed time penalty within 1 minute of the modelled 2.7.

Quarterly	**Refit the density surface and re-solve all routes.**	Three zones changed materially inside a single 12-week window while the total stayed flat (§10).	Refit completed within two weeks of each quarter end.
Ongoing	**Log driver overrides of the advisory route without challenge, and review them at each refit.**	Drivers hold real-time information the retrospective surface cannot (§15, R-06).	Override log maintained; reviewed quarterly.

1 8

Limitations, Bias & Ethics

What dispatch data cannot tell you, and what it must not be used for.

LIMITATION	EFFECT ON THE FINDINGS	MITIGATION APPLIED
Dispatch records measure reporting, not crime	A dispatch log records calls that were made. Areas with higher reporting propensity, more patrol presence or better phone coverage will appear denser regardless of underlying incidence. This is the single largest bias in all police-derived crime mapping.	Stated prominently; all findings framed as *recorded incident density*, never as crime rate. No comparison is drawn between areas as places.
Geocoding is generalised	Published locations are typically snapped to a block centroid or street segment, introducing a positional error on the order of tens to low hundreds of metres.	The 0.61 km kernel bandwidth is an order of magnitude larger than the expected geocoding error, so the surface is not sensitive to it.
Exposure is a proxy, not a prediction	Risk-kilometres measure how much high-density ground a vehicle crosses. They do not estimate the probability that this vehicle is victimised, which depends on target attractiveness, vehicle type and driver behaviour — none of which is in the corpus.	Every result stated as exposure reduction; no incident-prevention claim is made anywhere in this report (§14).
The router optimises over open ground	The lattice permits movement in eight directions between adjacent cells. Real vehicles are constrained by the road network, one-way systems, turn restrictions and weight and height limits. Modelled distances are therefore lower bounds.	Drivability check mandated before any route goes live (§17). Distance penalties should be read as indicative ratios rather than exact kilometres.
Retrospective surface, forward decision	The surface describes the 12-week window it was fitted on. Applying it to next month assumes stability, and §10 shows that assumption fails at zone level within a single window.	Quarterly refit required; three named zones flagged for review at every refit.
No control for exposure population	Density is per unit area, not per person or per vehicle passing. A busy commercial core has more incidents partly because it has more of everything.	Accepted deliberately: a routing decision is about the ground a vehicle crosses, so per-area density is the correct denominator for this use. It is the wrong denominator for any other question.
Synthetic locations in this issue	Every coordinate in this issue is generated. Zones Z-01 to Z-49 correspond to no real place.	Stated in §00 and repeated here. No figure may be cited as an observed measurement of any real location.

BINDING CONSTRAINTS ON USE

These follow from the limitations above and form part of the delivery terms.

1. The outputs may be used to route vehicles and to schedule stops. They may **not** be used to make any decision about a person — not employment, not insurance pricing, not credit, not tenancy, and not any assessment of an individual's risk.

2. Zone identifiers are deliberately neutral. Published or circulated versions of these maps must not be relabelled with neighbourhood names, because a density surface derived from reporting propensity relabelled as a place name becomes a claim about a community that this data cannot support.

3. The surface must not be presented to drivers as a prediction of where they will be attacked. It is an advisory cost layer, and rule R-06 preserves the driver's authority to override it.

4. Any published derivative must carry the reporting-bias statement from the first row of the table above.

STATISTICAL DISCLOSURE

No result was selected on the basis of significance. All 49 zones, six categories and four routes defined before analysis are reported whether or not they reached significance, including the non-significant citywide trend in §11 and the non-significant emerging zones in §10. The bandwidth sensitivity test in §07 was specified before the surface was fitted. The full incident table is delivered so every figure can be recomputed independently.

19

Methodology Appendix

Formula glossary and variable dictionary.

Formula glossary

QUANTITY	EXPRESSION	NOTES
Kernel density	$\hat{f}(\mathbf{s}) = (nh^2)^{-1} \Sigma K((\mathbf{s}-\mathbf{s}_i)/h)$	Bivariate Gaussian kernel; scaled by n to read as incidents per km².
Scott bandwidth	$h = n^{-1/(d+4)}\hat{\sigma}$	d = 2 dimensions; computed per surface.
Hotspot threshold	$\{\mathbf{s} : \hat{f}(\mathbf{s}) \geq Q_{0.95}\}$	95th percentile of the surface; minimum extent 0.12 km².
Concentration lift	(incident share) / (area share)	1.0 under a uniform distribution.
Gini of density	$(n+1-2\Sigma C_i/C_n)/n$	Inequality of the density surface across cells.
Zone risk index	$100(L_z-\min L)/(\max L-\min L)$	L = severity-weighted load per km²; relative to this study area only.
Nearest-neighbour index	$\bar{d}_{obs} / [2\sqrt{(n/A)}]^{-1}$	<1 clustered, =1 random, >1 dispersed.
Moran's I	$(n/\Sigma\Sigma w) \cdot (\Sigma\Sigma w_{ij} z_i z_j)/\Sigma z_i^2$	Queen contiguity; normality-assumption variance.
Split-half difference	$\Delta(\mathbf{s}) = \hat{f}_{late}(\mathbf{s}) - \hat{f}_{early}(\mathbf{s})$	Weeks 7–12 minus weeks 1–6, both night only.
Rate-ratio test	$B \sim \mathrm{Binomial}(a+b, 0.5)$	Conditional test on the split of counts between halves.
Edge cost	$c = \ell(1 + \lambda\hat{r})$	Mean normalised risk of the two endpoint cells.
Route exposure	$\Sigma\, \hat{r} \cdot \Delta\ell$	Line integral of normalised risk; units risk-km.
Shortest path	Dijkstra over the 8-connected lattice	19,600 nodes; solved once per λ.
Trend model	$y_t = \beta_0 + \beta_1 t + \varepsilon$	OLS on the weekly series, t = 1…12.

Variable dictionary — delivered incident table

FIELD	TYPE	DEFINITION
x_km	float	Easting in kilometres from the study-area origin.
y_km	float	Northing in kilometres from the study-area origin.
zone	categorical (49)	Reporting zone, Z-01 to Z-49.
category	categorical (6)	Incident type after the call-type filter.
week	integer 1–12	Collection week of the dispatch timestamp.
hour	integer 0–23	Local hour of the dispatch timestamp.
night	binary	1 where the hour falls in 22:00–04:59.
severity	float	Ordinal severity weight of the category.

2 0

Deliverables & Authorisation

What is handed over, and confirmation to proceed.

Delivered with this report

ITEM	CONTENTS
01 Structured database	Incident-level table, 85,556 rows × 8 fields, in XLSX and CSV, carrying coordinates, zone, category, week, hour, night flag and severity weight for every record.
02 Density surfaces	Day, night, combined and difference grids as arrays on the 140 × 140 lattice, ready to load as a routing cost layer.
03 Analysis script	Documented Python that regenerates every surface, statistic and route in this document from the delivered table — no manual step between the data and the page.
04 Figure repository	All 17 charts as vector SVG, named to their figure numbers, for reuse in operational briefings at any size.

Continuation options

OPTION	SCOPE
Live feed connection	The same pipeline pointed at a published open dispatch feed for the client's own city, producing a directly comparable issue of this report on observed data.
Quarterly refit	Rolling 12-week windows with hotspots, emerging zones and all routes re-solved each cycle, with change flagged against this baseline.
Routing integration	Delivery of the cost surface in the format your routing software consumes, with λ tuning and a drivability pass against the road network.
Fleet extension	Extension to the full destination set and to daytime windows, on the volume terms set out in the governing commercial proposal.

Commercial note

This report is issued under the terms of the Data Tune commercial proposal governing the engagement. Volume is measured in data entities on the same basis as that document; any extension beyond the contracted ceiling proceeds only on prior written approval, and never retrospectively.

Scheduled updates

Progress reporting during the working window follows the same three-point structure as the governing proposal: an early-stage note with initial counts and a first sample batch, a mid-project note with running totals, and a pre-delivery note with near-final counts before handover.

Acceptance & authorisation

By signing below, the client confirms receipt of this report and its accompanying deliverable set, accepts the scope and method recorded in §02 to §07, and acknowledges both the data provenance statement in §00 and the binding constraints on use in §18.

FOR THE CLIENT — AUTHORISED SIGNATURE

NAME & DESIGNATION · DATE

FOR DATA TUNE — AUTHORISED SIGNATURE

K. H. Militha Mihiranga

DATA SOLUTIONS CONSULTANT · DATE

Data Tune.

DT LINUX · DATA SOLUTIONS & DIGITAL CATALOGING

For clarification on scope, surfaces, index construction or scheduling, please contact the consultant alongside.

Prepared by	**K. H. Militha Mihiranga · Data Solutions Consultant**
Office	**555/24 Ranmuthugala, Kadawatha, Sri Lanka**
Email	**info@dtlinux.com**
Telephone	**+94 77 527 1186**
Website	**www.dtlinux.com**