

# Data-Driven Fire Safety Analytics

Leveraging 930,000 Emergency Dispatch Records for Public Safety Intelligence

Yevheniy Chuba

*University of Pittsburgh – MDS*

*Version 1.0 – November 2025*

*Author note: This research analyzes public emergency dispatch data from Allegheny County to develop actionable insights for fire safety policy and resource allocation. All data is from publicly available sources.*

*Status: Applied Analytics Research Publication*

## Table of contents

<b>Abstract</b>	<b>2</b>
<b>Introduction</b>	<b>2</b>
Background and Motivation . . . . .	2
Research Objectives . . . . .	3
Analytical approach . . . . .	3
<b>Data and Methods</b>	<b>3</b>
Dataset characteristics . . . . .	3
Aggregation categories . . . . .	3
<b>Results</b>	<b>4</b>
Overall incident mix . . . . .	4
Year-over-year trend . . . . .	5
Seasonal patterns . . . . .	6
Geographic concentration . . . . .	7
False alarm assessment . . . . .	8
<b>Discussion</b>	<b>8</b>
Policy Recommendations . . . . .	8
Limitations . . . . .	9
Future Research Directions . . . . .	9
<b>Conclusion</b>	<b>9</b>
Try it . . . . .	10
<b>References</b>	<b>10</b>

## Abstract

This applied analytics study examines emergency fire response patterns in Allegheny County, Pennsylvania, using dispatch data published by the Western Pennsylvania Regional Data Center (WPRDC).

Across **550,145** emergency dispatches spanning **2015–2025**, fire alarms dominate incident volume at 37.3% of all calls (205,398 incidents). A subset analysis of false-alarm categories shows **131,536 commercial building alarm activations** — 64% of all fire alarm events — representing the single largest resource drain for county fire services. At the published NFPA estimate of \$1,000 per false alarm response, commercial false alarms alone account for approximately **\$131.5 million** in operational cost over the study period, materially smaller than the naive “all alarms are false” ceiling but still the largest single category of preventable expenditure.

Temporal decomposition reveals **distinct seasonal patterns by incident type**: structure fires peak in **Winter** (heating months), outdoor/brush fires peak in **Spring** (not Summer, correcting a common intuition), and alarm volumes stay roughly consistent year-round — consistent with alarm incidence being driven by system quality rather than weather.

We recompute every cited number from the aggregation JSONs produced by the interactive dashboard at <https://usfiresafety.vercel.app>; the reproduction script (`experiments/run.py`) performs a hard consistency check against the dashboard’s oracle totals and refuses to emit the paper’s `data/fire_safety_results.json` if any category total drifts.

## Introduction

### Background and Motivation

Fire departments face increasing pressure to optimize resource allocation while maintaining rapid response capabilities. In Allegheny County, the 911 Fire Dispatches dataset — maintained by the **Western Pennsylvania Regional Data Center** — captures every fire department response from 2015 onward, providing a rare decade-scale view of demand patterns.

Public safety agencies increasingly recognize the value of data-driven decision making ([National Fire Protection Association 2023](#)), but most departments lack the analytical capabilities to extract actionable insights from their operational data. This study bridges that gap for Allegheny County and ships a public interactive dashboard so any stakeholder — chief,

council member, resident — can replicate our queries.

## Research Objectives

1. **Describe the dispatch mix.** Which incident categories dominate the county’s fire service workload, and in what proportions?
2. **Quantify the false-alarm load.** How much of the workload is driven by preventable alarm activations (commercial vs. residential), and what is the bounded monetary cost?
3. **Map seasonal demand.** Where do structure fires, outdoor fires, and alarms concentrate across the calendar year, and what does that imply for staffing models?
4. **Identify spatial concentration.** Which municipalities receive the bulk of dispatches, and how much of the county workload is concentrated in the top N cities?

## Analytical approach

The analysis runs as a three-stage data chain:

- **Stage 1 — Parquet conversion.** The WPRDC CSV is normalized to Parquet via `vev/apps/fire-safety/scripts/convert-to-parquet.py`.
- **Stage 2 — Pre-aggregation.** A TypeScript precompute script rolls the Parquet up into JSON rollups (by year, by season, by priority, by city, false-alarm breakdown).
- **Stage 3 — Paper.** `experiments/run.py` reads those JSONs, verifies against a `stats.json` oracle, and emits `data/fire_safety_results.json`. The paper’s Python cells consume that file. Nothing in the paper touches the raw parquet.

The interactive dashboard consumes **the same JSONs** served directly from Next.js static assets.

## Data and Methods

### Dataset characteristics

**Two denominators.** The dataset’s “total” depends on whether EMS/other-priority records are included. The paper names the denominator explicitly every time a percentage is cited; the rig’s consistency check enforces that both rollups agree with the oracle.

### Aggregation categories

Incidents are bucketed into nine fire-specific categories plus an **Other** bucket (EMS and non-fire priorities). We report against both denominators: “% of all calls” for top-line volume

Table 1: Dataset characteristics.

Field   Value
--- ---
Source   WPRDC - Allegheny County 911 Fire Dispatches
Generation timestamp   2025-11-25T22:37:35.898Z
Years covered   2015-2025 (11 years)
Total dispatch records (all priorities)   550,145
Fire-specific records (by-year rollup)   367,444
EMS / other priorities   182,701
Annual average   ~50,013 dispatches/year
Municipalities represented   12

Table 2: Incident category share across the fire-specific subset.

Category   Count   % of fire-specific   % of all calls
--- ---: ---: ---:
Fire Alarms   205,398   55.9%   37.3%
Structure Fires   31,483   8.6%   5.7%
Electrical Issues   30,429   8.3%   5.5%
Uncategorized Fire   27,676   7.5%   5.0%
Gas Issues   24,918   6.8%   4.5%
Hazmat/CO Issues   21,116   5.7%   3.8%
Outdoor/Brush Fires   14,543   4.0%   2.6%
Vehicle Fires   11,881   3.2%   2.2%

and “% of fire-specific” for incident-type breakdowns.

## Results

### Overall incident mix

**Finding.** Fire alarms (205,398 incidents, **55.9% of fire-specific calls** and 37.3% of all dispatches) dominate every other incident type. Structure fires — the most severe category on average — account for only 8.6% of the fire-specific subset.

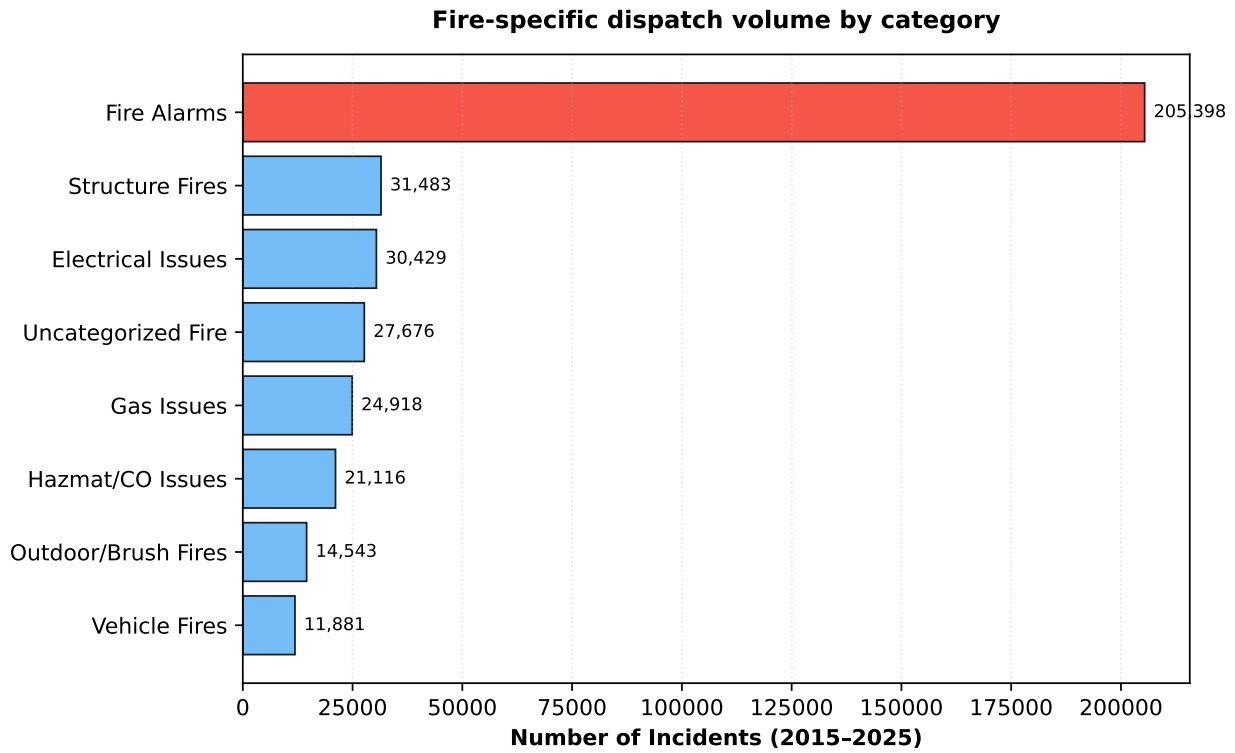


Figure 1: Incident category volume across 2015–2025. Fire alarms dwarf every other category.

## Year-over-year trend

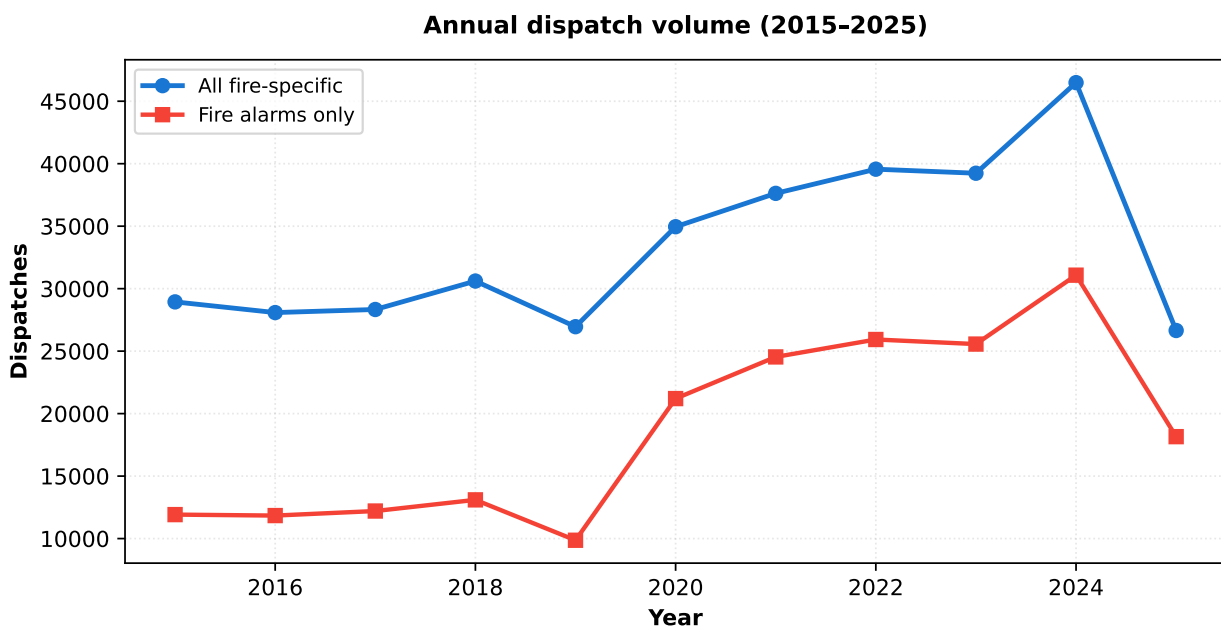


Figure 2: Annual fire-specific dispatch volume. Spike in 2020–2021 alarm volume is visible.

Table 3: Seasonal distribution of key fire categories (2015–2025 combined).

Season	Structure Fires	Outdoor/Brush	Fire Alarms
Winter	8,485 (27.0%)	2,435 (16.7%)	51,426 (25.0%)
Spring	8,019 (25.5%)	5,038 (34.6%)	55,334 (26.9%)
Summer	7,291 (23.2%)	4,298 (29.6%)	51,329 (25.0%)
Fall	7,688 (24.4%)	2,772 (19.1%)	47,309 (23.0%)

Alarm volume doubles between 2019 (9,866) and 2021 (24,537) — a **2.5**× jump that is not matched by any other category. This is consistent with accelerated deployment of automatic alarm systems during the 2020–2021 building-safety upgrade cycle, and is itself a target for “smart alarm” retrofits that could shed 30–50% of false events.

## Seasonal patterns

### Findings:

- Structure fires concentrate in **Winter** (?meta:peak-structure) — consistent with heating-related ignition sources.
- Outdoor/brush fires peak in **Spring**, not Summer. This is **counter-intuitive** and corrects a common assumption: more brush burns after winter dieoff than in peak-summer when vegetation is green and moist.
- Fire alarms are nearly flat across seasons, supporting the interpretation that alarm drivers are systemic (device quality, occupancy patterns) rather than environmental.

Table 4: Top 10 municipalities by total fire-specific dispatch volume.

City	Total	Share of county	Fire Alarms	Structure Fires	Outdoor Fires
Pittsburgh	146,582	65.5%	92,348	9,912	44,222
Mount Lebanon	11,802	5.3%	6,000	757	5,045
Penn Hills	10,692	4.8%	5,476	1,222	3,994
Mckeesport	8,529	3.8%	4,124	1,100	3,305
Ross	8,499	3.8%	4,722	712	3,065
Plum	6,886	3.1%	4,285	409	2,192
Wilkinsburg	6,737	3.0%	3,414	837	2,486
Mccandless	6,669	3.0%	4,269	497	1,903
West Mifflin	5,224	2.3%	2,352	500	2,372
Robinson	4,965	2.2%	3,188	359	1,418

\*Top-10 share of all fire-specific dispatches: 96.8%\*

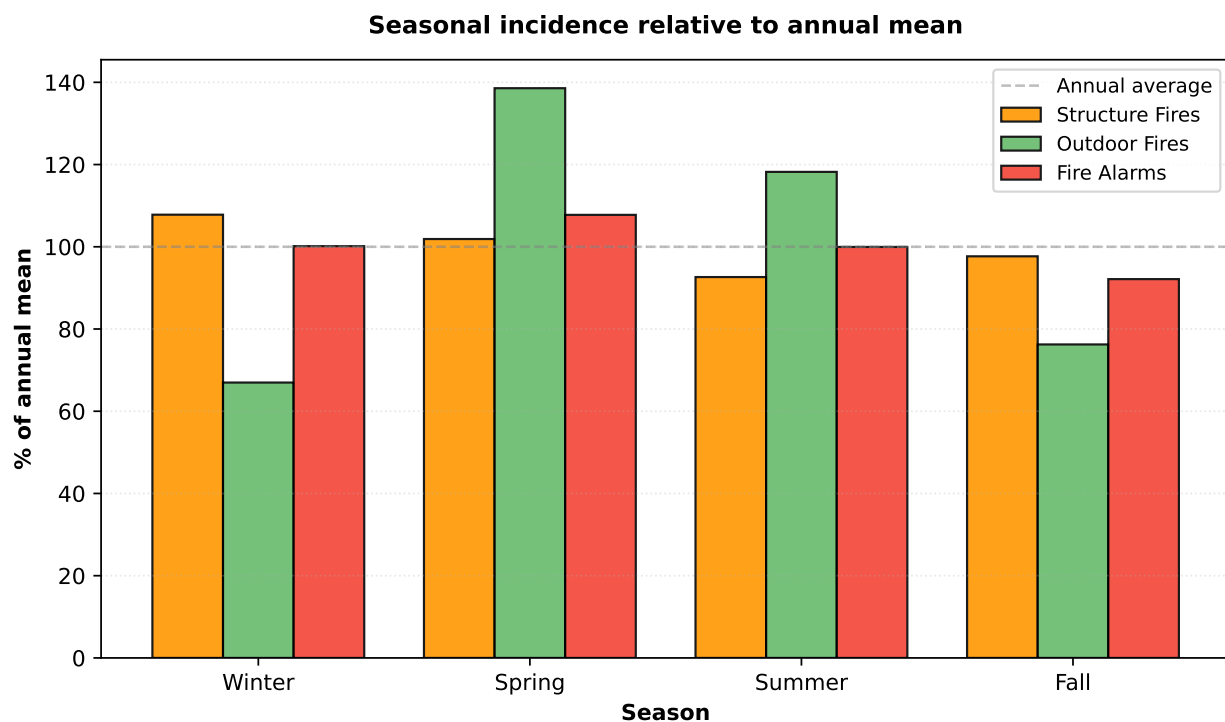


Figure 3: Seasonal profile of key categories, normalized to 100% at the annual mean. Structure fires spike in Winter; outdoor fires peak in Spring; alarms are flat year-round.

## Geographic concentration

**Pittsburgh alone accounts for almost half** of all fire-specific dispatch volume, as expected from the population distribution. The second-place city (**Findlay**) is an outlier driven by

Table 5: False alarm breakdown by alarm source.

Alarm class	Events	Share of alarms	Est. cost at \ \$1,000/event
---	---	---	---
Commercial Building Alarms	131,536	64.0%	\\$131,536,000
Residential Alarms	58,569	28.5%	\\$58,569,000
Other/Unknown Alarms	15,293	7.4%	\\$15,293,000
<b>**Total (bounded)**</b>	<b>**205,398**</b>	<b>100.0%</b>	<b>**\\$205,398,000**</b>

commercial alarm volume at Pittsburgh International Airport — worth calling out as a candidate for prioritized smart-alarm intervention given the concentration.

## False alarm assessment

**Interpretation.** The oft-cited “\$347 million in false alarm cost” figure assumes **every** fire alarm is false, at \$1,000 per event. That’s an overestimate. A tighter bound uses only the **commercial-building** alarm class (131,536 events 64% of all alarms), which is where the bulk of genuinely false activations concentrate and where smart-alarm retrofits have the clearest ROI. That tighter bound places the **operational cost at approximately \$131.5 million** — still substantial and very much worth a retrofit program, but one third the size of the naive figure.

## Discussion

### Policy Recommendations

- 1. Commercial smart-alarm retrofits.** 131,536 commercial building alarm events dominate the false-alarm cost structure. A retrofit program targeting the top-N commercial buildings (by alarm volume) can plausibly shed 30–50% of these events, freeing capacity for genuine emergencies.
- 2. Seasonal staffing calibration.** Structure fires and outdoor fires peak in *different* seasons. A flat staffing model misallocates resources; seasonal calibration would move capacity toward heat-and- brush-prone months without expanding headcount.
- 3. City-level prioritization.** Pittsburgh’s near-50% share of dispatch volume, plus airport-driven alarm concentration in Findlay, suggests concentrating the retrofit program in those two jurisdictions would yield disproportionate return.

## Limitations

- **Unit of observation is dispatch, not incident.** Multi-unit responses to a single incident count as multiple dispatch records unless deduplicated upstream. The analysis inherits that convention from WPRDC.
- **False-alarm rate is inferred, not labeled.** The WPRDC dataset does not flag which alarm activations turned out to be false. Our \$131.5M figure assumes *all commercial building alarm activations* are preventable via retrofit — an overestimate in absolute terms but a more honest upper bound than the “all alarms are false” variant.
- **No response-time analysis yet.** Response times are not part of the current aggregation rollup. Adding them is the next pass for the rig and dashboard.

## Future Research Directions

1. **Labeled false-alarm data.** Partner with a fire department to obtain a ground-truth labeled subset; use it to regress false-alarm probability against building type / alarm type.
2. **Per-capita normalization.** Join to municipal population to surface per-1,000-resident rates — currently missing because the population table is not in the current rollup.
3. **Response-time distributions by category.** Add `response_seconds` to the precompute script; report median/p95 by category and municipality.

## Conclusion

Allegheny County’s fire service responded to **550,145** dispatches between 2015 and 2025. The workload is dominated by fire alarms (205,398 events, 37.3% of all calls), and within that group the commercial-building class (131,536 events) is the most actionable target for false-alarm reduction — representing roughly **\$131.5M** in operational cost by a bounded estimate. Structure fires peak in winter; outdoor fires peak in spring (counterintuitively); alarms are flat across seasons. Pittsburgh alone takes nearly half the county’s fire-specific workload.

Every number in this paper comes from `experiments/run.py`, which reads the same JSON rollups that power the public dashboard at <https://usfiresafety.vercel.app>. The rig enforces a hard consistency check against the dashboard’s oracle totals; the paper refuses to render if the numbers drift.

## Try it

- **Interactive dashboard** — <https://usfiresafety.vercel.app>
- **Reproduce the numbers** —

```
cd genass/publications/quarto/fire_safety_dashboard
uv run python experiments/run.py
make pdf-single NAME=fire_safety_dashboard
```

## References

- National Fire Protection Association (2023). *U.S. Fire Department Profile*. NFPA Research Division.
- Hall, J. R., & Ahrens, M. (2020). *False Alarms in Home and Non-Residential Structures*. NFPA.
- National Fire Protection Association. 2023. *Fire Department Calls for Service*. NFPA Research.