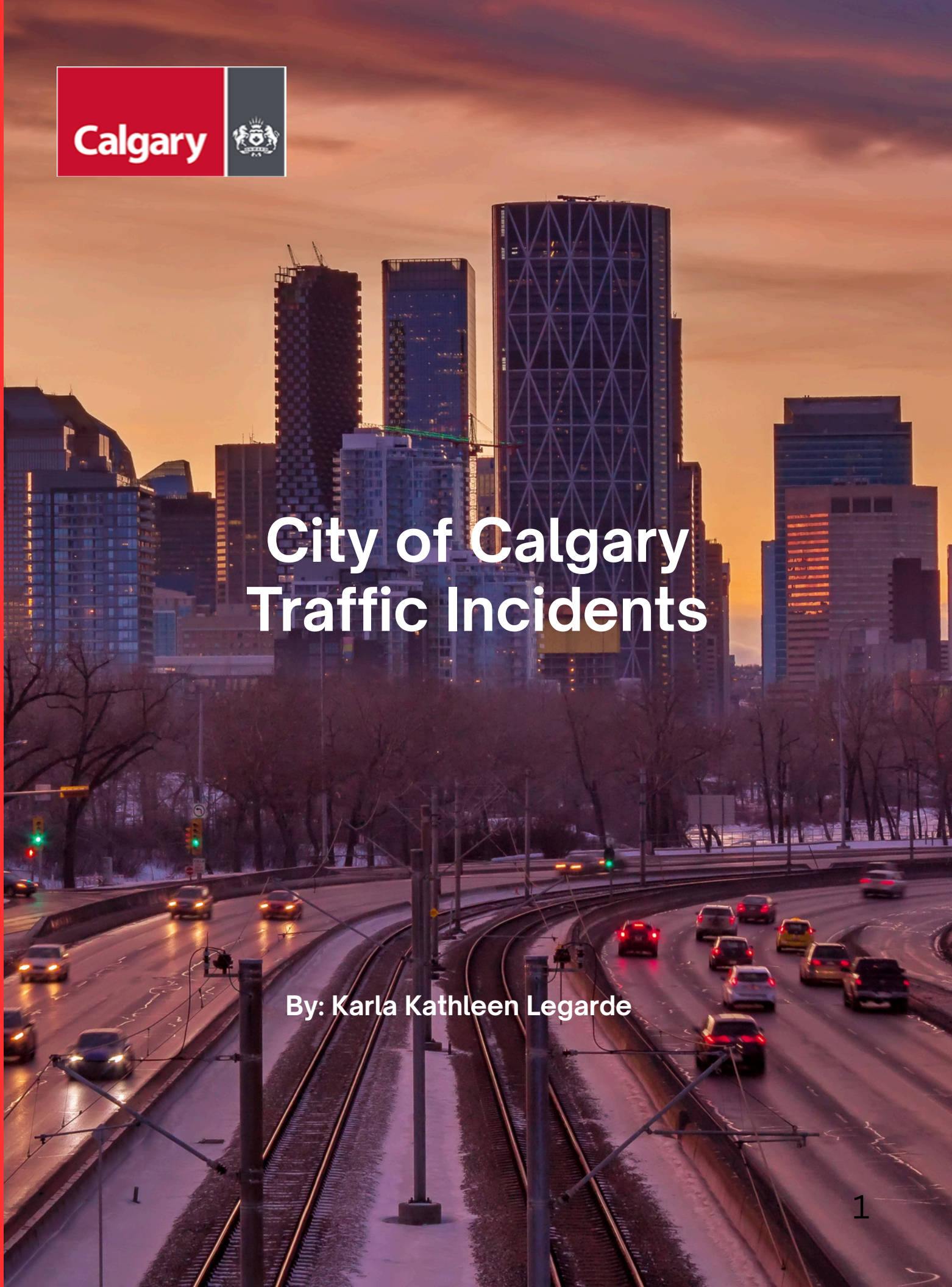


City of Calgary Traffic Incidents

By: Karla Kathleen Legarde

CONTENT

INTRODUCTION

Traffic incidents are an inevitable part of urban life, impacting safety, mobility, and city operations. For a growing city like Calgary, understanding when and where these incidents occur is essential to improving road safety and managing infrastructure effectively.

This project analyzes open-source data from the City of Calgary Open Data Portal, focusing on reported traffic incidents between January 2022 and May 2025. Using Microsoft Excel for data preparation and Power BI for modeling and visualization, over 25,000 records were transformed into an interactive analytical model.

The goal of the study is to uncover temporal and spatial patterns in traffic incidents, by time of day, day of week, and location, and to distinguish between accident-related and non-accident events. These insights aim to support public management initiatives by highlighting high-risk zones, identifying peak incident periods, and informing future strategies for safer and more efficient urban mobility.

PROJECT GOAL

The primary purpose of this project is to analyze traffic incident trends in Calgary to uncover actionable insights about time, location, and type of incidents.

By identifying recurring patterns, the project aims to support city management goals such as safety enhancement, infrastructure improvement, and efficient resource allocation.

Specific objectives include:

1. To explore and visualize when and where traffic incidents most frequently occur.
2. To distinguish between accident-related and non-accident-related incidents.
3. To evaluate year-over-year changes in traffic incident volume.
4. To support public awareness and strategic planning using data-driven evidence.

While the project is an academic and civic data initiative rather than an official city program, the insights derived can contribute to Calgary's broader goals of safety and mobility optimization. It also serves as a strong portfolio example of applying analytics for public good.

ABOUT THE DATASET

The dataset used for this project, titled “Traffic Incidents,” is publicly available through the City of Calgary Open Data Portal.

It is an open government dataset that documents all traffic-related incidents reported to the City, offering a transparent perspective on road safety and mobility within Calgary.

The dataset contains several key categories of information, including:

- Incident Information: unique identifiers, descriptive details, event status, and type of incident
- Temporal Data: start and end date/time, as well as update timestamps
- Location Data: latitude, longitude, quadrant (NW, NE, SW, SE), and address details
- Impact Type: severity level and, when available, the number of affected lanes

For the purpose of this analysis, the dataset was filtered to include records from January 1, 2022, to May 20, 2025, resulting in approximately 25,000 cleaned incident records. This time frame was intentionally selected to capture post-pandemic traffic patterns and recent urban developments that may have influenced incident distribution.

The dataset’s granular structure where each record represents a single reported incident enabled both temporal and spatial exploration.

However, in its raw form, the dataset exhibited several data quality and structural issues that required thorough cleaning and transformation before it could be effectively modeled and analyzed.

KEY QUESTIONS

The analysis was guided by several key questions designed to uncover meaningful trends in Calgary's traffic incidents and support data-driven insights for urban mobility and safety. Each question is linked to a specific analytical objective and corresponding visual within the Power BI dashboard.

1. What types of incidents occur most frequently in Calgary?
2. When do traffic incidents most often occur — by hour, day, or month?
3. Which city quadrants or areas experience the highest number of incidents?
4. How have total incidents changed over time (year-over-year)?
5. How frequent are pedestrian incidents in Calgary's downtown and high-traffic zones?

Together, these questions provided a structured analytical framework that connected descriptive statistics with spatial and temporal trends, allowing for deeper understanding of Calgary's traffic dynamics.

METRICS AND KEY PERFORMANCE INDICATORS

To evaluate traffic incident patterns and monitor performance across time, several Key Performance Indicators (KPIs) were defined. Each KPI supports the project's strategic and operational goals by quantifying different aspects of incident frequency, type, and location.

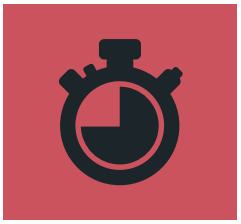
Overall Executive

Growth
Quadrants (Areas)
Periods

Regions Where

Distribution
(Key Quadrants)
Density

Time When



Daily (Hours)
Weekly (Days)
Monthly

Pedestrian Areas

Accidents related to pedestrians
Frequency
(Hours)

Each KPI supports the analytical goals of the project by transforming raw incident records into measurable insights. Together, they form the foundation for visual storytelling within the Executive Dashboard and Analytical Dashboard, helping to identify high-risk periods, monitor safety trends, and guide data-informed decision-making.

DATA CLEANING AND TRANSFORMATION

Before analysis and visualization were conducted, the raw dataset underwent an extensive data cleaning and transformation process to ensure accuracy, consistency, and readiness for modeling. All data preparation and structuring were performed in Microsoft Excel, where the dataset was refined and reshaped into multiple tables following a dimensional design.

1. Initial Data Cleaning

The initial phase focused on improving data accuracy and consistency.

- Irrelevant columns that did not contribute to the analysis were removed to simplify the dataset and enhance processing efficiency.
- Column headers were standardized to improve readability and maintain uniform naming conventions (e.g., START_DT renamed to StartDate).
- Data types were corrected to match field content (dates were converted to Date/Time formats), and location fields were stored as numeric values.
- Text formatting was standardized by trimming whitespace and correcting inconsistent capitalization across descriptive columns.
- Missing values in the “Quadrant” field were properly handled by reassigning quadrants based on latitude and longitude coordinates, ensuring that every record was geographically classified under one of Calgary’s four official quadrants (NW, NE, SW, SE).

This step ensured that the dataset was free of redundant and inconsistent entries before further transformation.

2. Creation of Dimension Tables

Once the master dataset was cleaned, dimension tables were created in Excel to support a star schema design and enable efficient analysis once loaded into Power BI. The process involved:

- Extracting unique values and creating lookup tables for key entities such as Date, Time, Location, and Incident Type.
- Assigning unique identifiers (keys) to each dimension table to facilitate relational connections within Power BI.
- Designing the Fact_Incident table as the central table, containing all incident records with foreign keys linking to each dimension.
- Each table was placed on a separate worksheet to mirror a relational database structure, allowing for efficient linking once imported into Power BI.

This transformation ensured that the dataset followed a structured and scalable model, optimized for filtering, aggregation, and visualization in Power BI.

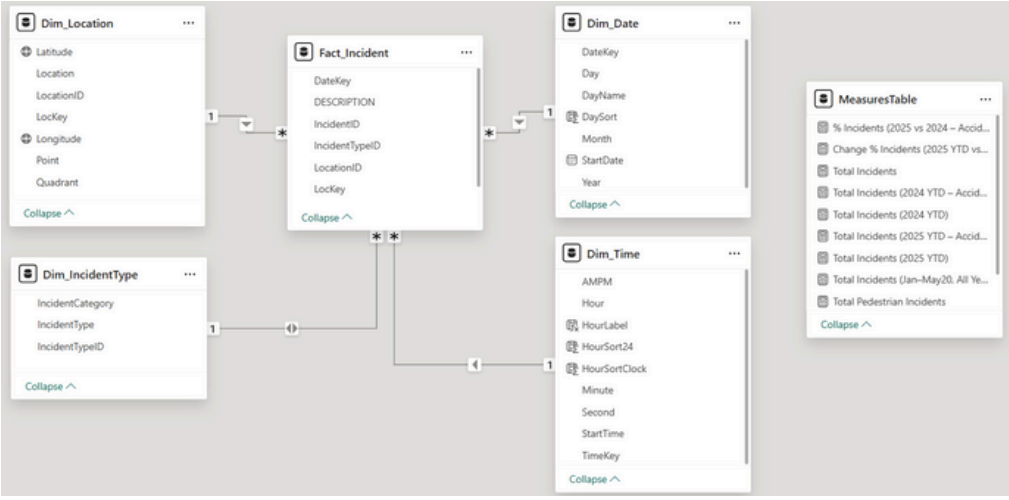
3. Validation and Final Preparation

Once the Fact and Dimension sheets were completed, several validation checks were performed to ensure readiness for visualization:

- Row counts were verified to confirm that all original incidents were preserved after cleaning and transformation.
- Primary and foreign keys were tested in Excel to ensure that each relationship (e.g., DateKey, TimeKey, LocationKey) matched correctly.
- Quadrant recalculations were cross-checked using coordinate samples to confirm geographic accuracy.
- Date continuity was verified to ensure coverage for the entire analysis period from January 2022 to May 2025.

The result was a fully structured, validated, and analysis-ready dataset composed of a Fact table and four Dimension tables, each in its own Excel worksheet. This format provided a solid foundation for import into Power BI, where relational modeling, measures, and visualizations were later developed.

DATA MODEL



After cleaning and restructuring the dataset in Excel, the prepared tables were imported into Power BI Desktop for data modeling and visualization. Within Power BI, a star schema was established to enable efficient analysis, relationship management, and the creation of interactive dashboards.

The data model consisted of a central Fact table supported by four Dimension tables, each representing a specific analytical perspective: date, time, location, and incident type. A separate Measures Table was also created to store calculated DAX measures used across reports.

Relationship Design

Relationships were created using Power BI's Model View, with clear cardinalities and single-direction filters to maintain data integrity.

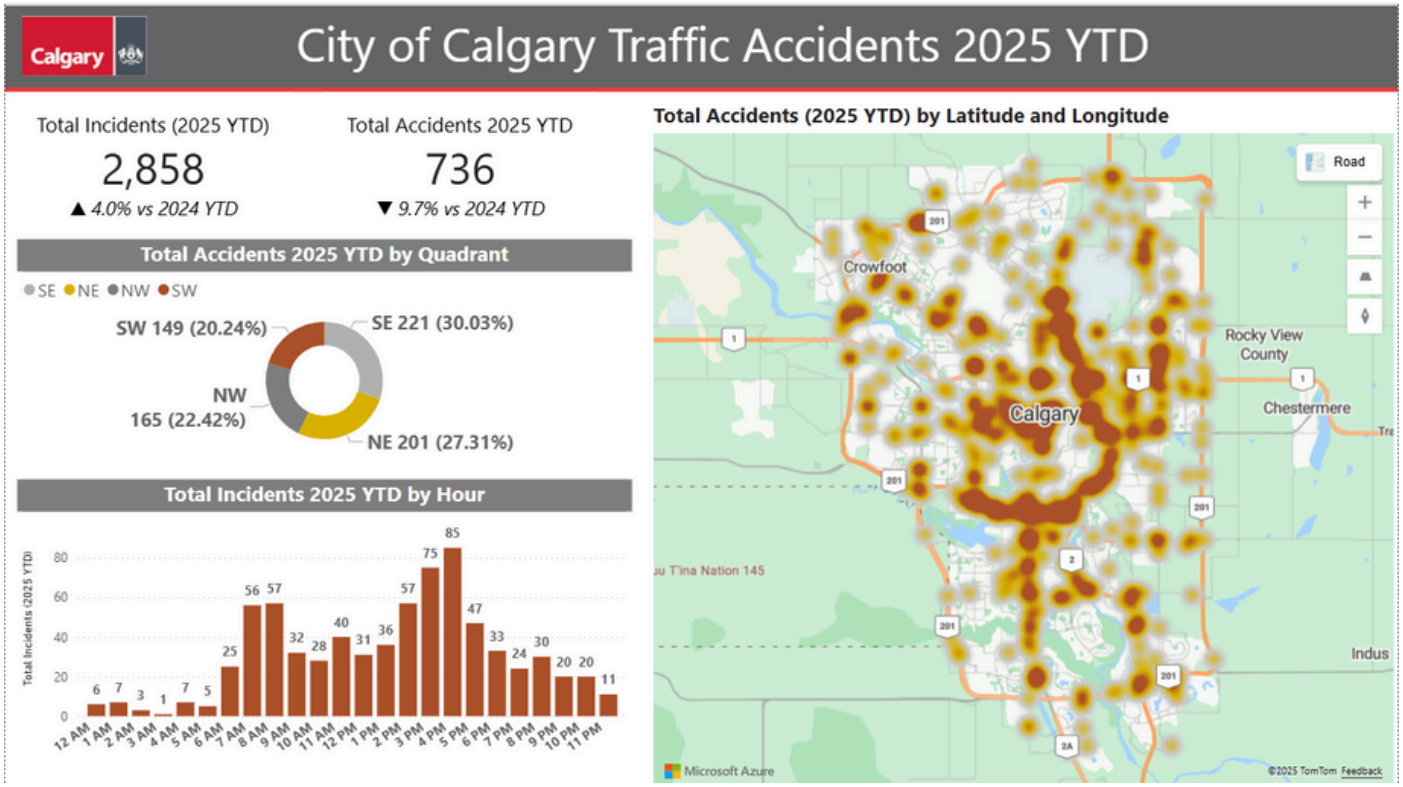
- Dim_Date[DateKey] → Fact_Incident[DateKey]
- Dim_Time[TimeKey] → Fact_Incident[TimeKey]
- Dim_Location[LocationID] → Fact_Incident[LocationID]
- Dim_IncidentType[IncidentTypeID] → Fact_Incident[IncidentTypeID]

All relationships followed a one-to-many (1:*) structure, where each dimension table (1) filters the central fact table (*). This approach aligns with best practices in dimensional modeling and ensures optimal query performance in Power BI.

RESULTS

The Power BI dashboards developed for this project provided both a high-level overview and detailed insights into Calgary's traffic incident trends from January 2022 to May 2025. Four dashboards were created, each addressing a different aspect of analysis and stakeholder need.

EXECUTIVE DASHBOARD



The Executive Dashboard provides a snapshot of Calgary's 2025 year-to-date traffic conditions, highlighting total incidents, accident distribution, time patterns, and spatial concentrations across the city.

Headline KPIs:

- 2,858 total incidents (▲ 4.0 % vs 2024) and 736 total accidents (▼ 9.7 % vs 2024) indicate that overall incident volume rose slightly, but accidents declined, suggesting gradual safety improvement.

Quadrant Distribution:

- The Southeast (30%) and Northeast (27%) recorded the most accidents, followed by the Northwest (22%) and Southwest (20%), consistent with heavier commuter activity in eastern corridors.

Hourly Pattern:

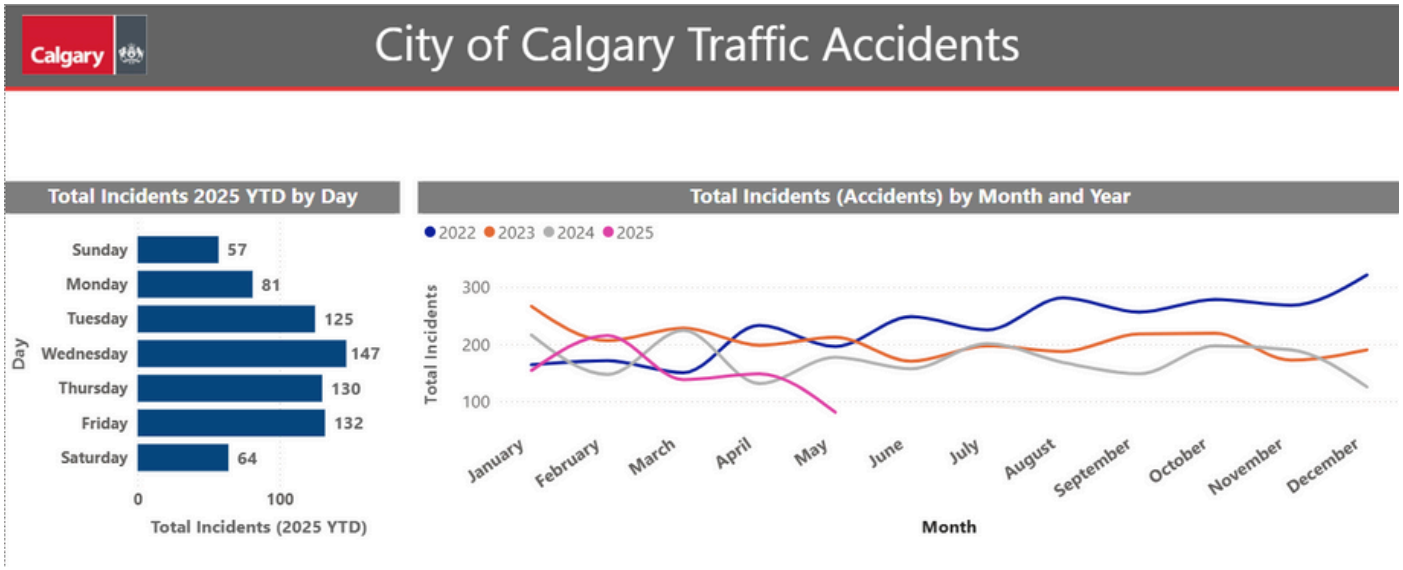
- Accidents rise during morning traffic (5–7 AM) and peak in the late afternoon (3–5 PM), confirming rush hour as the most accident-prone period.

Spatial Concentration:

- Heatmap clusters appear along Deerfoot Trail and Glenmore Trail with dense activity around NE and SE, reflecting persistent high-traffic risk zones.

Interpretation:

Despite fewer accidents overall, concentration in the SE and NE quadrants and recurring afternoon peaks suggest continued monitoring and targeted safety measures along Calgary's busiest corridors.



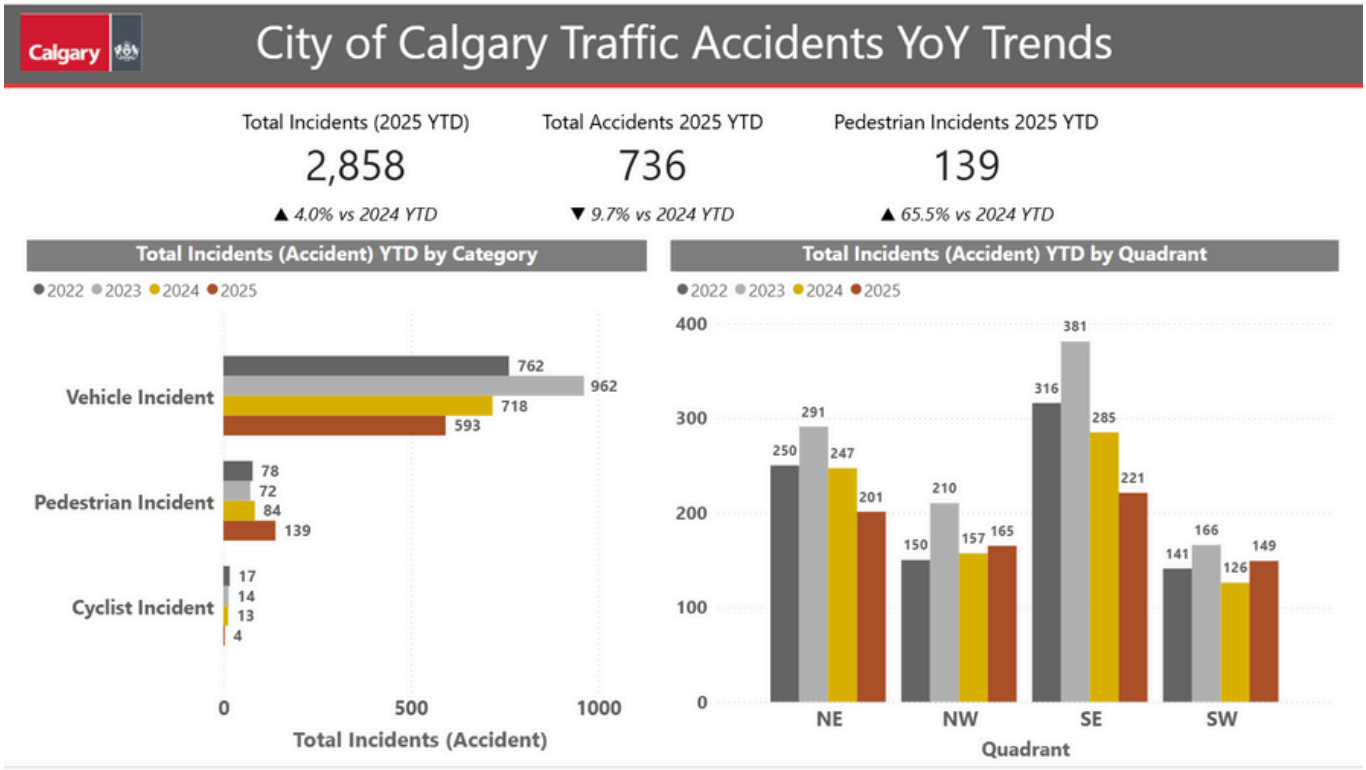
This dashboard summarizes the total accident distribution by weekday and compares incident volumes across months and years (2022–2025).

- **Weekday Pattern:** Wednesday recorded the highest number of accidents (147), followed by Tuesday (125) and Thursday (130). Sundays and Saturdays had the fewest, with 57 and 64 accidents respectively. This pattern reflects Calgary's commuter-driven weekday traffic volumes.

- **Monthly Trend:** The line chart shows that 2022 consistently experienced higher monthly totals compared with 2023–2025. The 2025 line (yellow) remained below the 2022 and 2023 levels, indicating a continued overall decline in accident frequency.

- The steady weekday peaks and year-over-year decline suggest improved traffic management and possibly reduced congestion since the pandemic years.

YEAR-OVER-YEAR TRENDS DASHBOARD

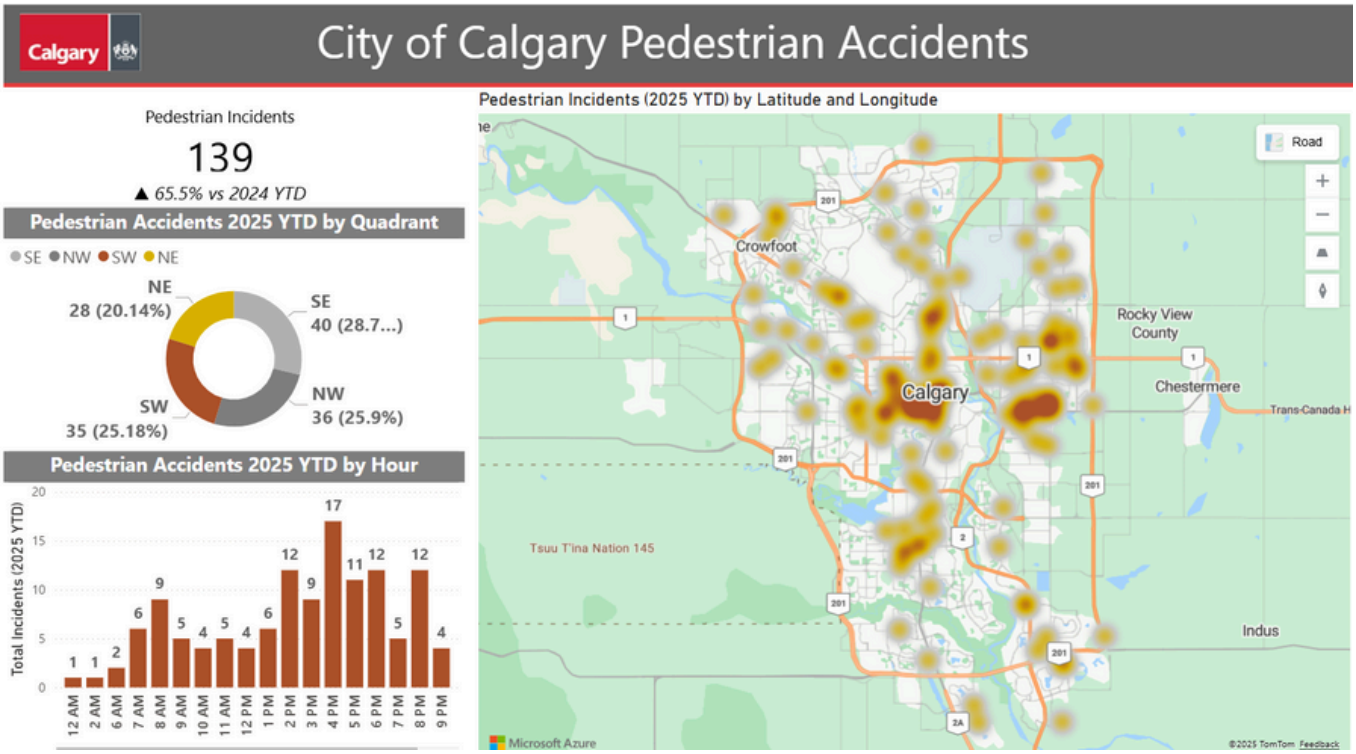


This dashboard highlights comparative performance indicators, incident types, and quadrant-level analysis.

- **Headline KPIs:**
  - Total Incidents (2025 YTD): 2,858, representing a 4.0 % increase from 2024 YTD.
  - Total Accidents (2025 YTD): 736, showing a 9.7 % decrease from 2024 YTD.
  - Pedestrian Incidents (2025 YTD): 139, marking a substantial 65.5 % increase from the previous year.
- **By Category:**
  - Vehicle collisions remain dominant (962 in 2022 → 593 in 2025).
  - Pedestrian accidents rose from 84 in 2024 to 139 in 2025, confirming a recent upward trend.
  - Cyclist incidents continued to decline (from 17 in 2022 to 4 in 2025).
- **By Quadrant:**
  - The Southeast (SE) quadrant consistently recorded the highest accident counts (381 in 2023 → 221 in 2025).
  - The Northeast (NE) also showed high concentration, while Southwest (SW) and Northwest (NW) experienced relatively fewer incidents.

Overall, while total reported incidents increased slightly, the proportion of serious accidents declined, indicating improved safety conditions for motorists but emerging risk areas for pedestrians.

PEDESTRIAN ACCIDENTS



Focusing exclusively on pedestrian-related incidents, this dashboard combines geographic and temporal perspectives.

- **2025 YTD Total:** 139 pedestrian incidents, a 65.5 % rise compared with 2024 YTD.
- **By Quadrant:** The Southeast (SE) quadrant accounted for the largest share (28.7 %), followed by the Northwest (25.9 %) and Southwest (25.2 %).
- **Time-of-Day Pattern:** Incidents peaked at 4 PM (17 cases), corresponding with afternoon rush-hour pedestrian traffic.
- **Spatial Distribution:** The heatmap showed dense clusters of pedestrian incidents around Downtown Calgary, particularly the Beltline and 17 Avenue SW, areas with high foot traffic and mixed-use development. Additional concentrations were also observed in the Forest Lawn area of the Southeast quadrant, indicating that pedestrian safety concerns extend beyond the city core into busy suburban corridors.

These findings underscore the growing safety concern for pedestrians in central business and entertainment districts, suggesting a potential need for enhanced crosswalk visibility, signal timing adjustments, or public awareness campaigns.

CONCLUSION

This analysis of the City of Calgary Traffic Incidents dataset demonstrates how open-source municipal data can generate valuable insights when properly cleaned, modeled, and visualized.

Using over 25,000 records (2022–2025), the study revealed clear spatial and temporal trends across Calgary's road network.

Results showed that while overall incidents slightly increased in 2025, accidents declined, suggesting gradual safety improvements.

However, pedestrian incidents rose sharply, particularly in Downtown Calgary and Forest Lawn (SE), highlighting the need for enhanced crosswalk design and public safety measures.

Accidents were most common during weekday rush hours (3–5 PM) and in the Southeast and Northeast quadrants, where major corridors such as Deerfoot Trail, Glenmore Trail, and Macleod Trail experience heavy traffic volumes.

These findings indicate that while roadway safety continues to improve, pedestrian protection and intersection management remain key priorities.

The project also reinforces the potential of open data and Power BI visualization to support data-driven decisions for safer, smarter urban mobility.