CALGARY AIR QUALITY ANALYSIS FOR 2024

DETAILED ANALYSIS REPORT
BY
KARAN ACHARYA

Executive Summary

This case study presents a comprehensive, year-long analysis of Calgary's air quality trends for the year 2024, utilizing open-source environmental data from **Alberta Environment & Protected Areas** and leveraging **Power BI** for advanced data visualization. The project is centred on assessing daily **Air Quality Index (AQI)** values and identifying the major pollutants affecting urban air health, with a goal of translating raw environmental data into actionable insights through interactive dashboards.

The dataset, comprising over 365 daily records, includes critical pollutants such as **PM2.5**, **PM10**, **NO_2**, **CO**, and **O_3**, alongside AQI classifications and category metadata. These values were analyzed in combination with spatial observations from **Sentinel-5P's TROPOMI instrument**, providing additional context on pollutant behavior at a higher atmospheric level. The study incorporates both numerical analysis and geospatial awareness, focusing on understanding pollutant concentration levels, seasonal fluctuations, and pollutant contributions to poor air quality conditions.

Key performance indicators (KPIs) such as **average AQI**, **maximum and minimum AQI values**, and the **distribution of good vs. poor AQI days** were used to evaluate air quality trends at both operational and strategic levels. The project also explored seasonal patterns and sought to answer core environmental questions, such as identifying the most frequent pollutants and understanding the monthly progression of AQI levels.

Notably, the findings indicate that **only 13% of the year experienced 'Good' air quality**, with the rest categorized as **Moderate or Poor**, primarily driven by elevated **PM2.5** and **PM10** concentrations. A significant AQI spike was observed in **September**, suggesting a combination of industrial activity, temperature inversion, or regional wildfire impact. Seasonal analysis revealed clear patterns, reinforcing the importance of data-driven intervention strategies.

By combining technical tools with open government data, this study not only offers a snapshot of Calgary's environmental health but also exemplifies how **data visualization platforms like Power BI can be effectively used for environmental monitoring, policy planning, and public awareness**. The case study advocates for continuous data collection, public transparency, and the development of predictive models to proactively manage air quality challenges in urban environments.

1. Information on the Open-Source Data of Interest

The primary data source used in this analysis is the **Alberta Open Government Portal**, which provides open-access environmental data. Specifically, we used a **CSV dataset containing daily Air Quality Index (AQI) readings for Calgary throughout the year 2024**.

Key Features of the Dataset:

- **Source Authority:** Alberta Environment & Protected Areas
- **Format:** Comma-Separated Values (CSV)
- **Location:** Calgary, Alberta, Canada
- **Timeframe:** January 1, 2024 – December 31, 2024
- **Observations:** Daily data entries for a full year
- **Key Columns:**
    - record_id: Unique identifier for each entry
    - location: Geographical location (Calgary)
    - AQI: Air Quality Index (quantitative value representing overall air quality)
    - PM2.5: Particulate matter ≤ 2.5 micrometers
    - PM10: Particulate matter ≤ 10 micrometers
    - NO_2: Nitrogen Dioxide concentration
    - CO: Carbon Monoxide concentration
    - O_3: Ozone concentration
    - Categories: Qualitative category assigned to AQI value (e.g., Good, Moderate, Poor)
    - category_id: Categorical code representing AQI classification

Supplementary Data Source:

- **Satellite Data:** The dataset insights are enriched with satellite monitoring from **Sentinel-5P**, a satellite operated by the European Space Agency (ESA), equipped with the **TROPOMI instrument** (Tropospheric Monitoring Instrument).
    - **Pollutants Monitored:** NO_2, CO, O_3, CH_4, SO_2, and aerosols
    - **Spatial Resolution:** ~5.5 km^2 per pixel
    - **Coverage:** Global (daily revisit over Calgary included)

This combination of on-ground and satellite-based open data ensures comprehensive spatial and temporal accuracy in tracking air quality trends.

2. Purpose and Goals

Primary Objective

The overarching purpose of this project is to conduct a comprehensive analysis of **Calgary's air quality trends for the year 2024** using open-source environmental data and advanced data visualization tools. The aim is to generate **data-driven insights** that can inform **environmental policy**, support **public health awareness**, and contribute to **academic and scientific discourse** on urban air quality management.

Specific Goals

This initiative aligns with the broader mission of **Environment and Climate Change Canada (ECCC)**, which emphasizes the importance of:

- Monitoring air quality trends across Canadian urban centers
- Evaluating the effectiveness of environmental regulations
- Identifying emerging environmental health risks
- Promoting sustainable development through informed policy

To achieve the primary objective, the project is structured around three specific goals:

1. Trend Analysis: Understanding Temporal Patterns

The first goal is to **uncover how air quality evolves over time**, including daily, monthly, and seasonal variations throughout 2024. Temporal analysis focuses on:

- Detecting **long-term trends** in overall AQI values
- Identifying **seasonal spikes** in pollutant levels (e.g., due to wildfire smoke, industrial activity, or weather phenomena)
- Understanding how certain **months or periods** (such as September) consistently show higher pollution levels

By breaking down the dataset into time-based intervals and visualizing trends using **line charts and time series graphs**, this goal provides insight into **when Calgary experiences its cleanest and most polluted air**, enabling **timely interventions** and **seasonal health planning**.

2. Pollutant Assessment: Analyzing Pollution Composition

The second goal is to **examine the composition of air pollution**, with a focus on identifying which pollutants contribute most significantly to AQI deterioration. Particular attention is given to:

- **Particulate Matter (PM2.5 and PM10)**, known for their harmful health impacts and prevalence in urban environments

- **Nitrogen Dioxide (NO_2)** and **Ozone (O_3)**, which are known to interact in sunlight to form smog
- **Carbon Monoxide (CO)**, typically associated with traffic emissions and heating systems

3. Public Health Communication: Simplifying Complex Data

Air quality data, while rich and informative, can be **technically complex** and difficult for the general public to interpret. Thus, a key goal of the project is to **convert raw environmental data into easy-to-understand, actionable visual insights**. This involves:

- Using **Power BI dashboards** with intuitive visuals such as:
  - KPI cards for instant snapshots
  - Pie charts for AQI category distribution
  - Bar and line graphs for pollutant trends and temporal shifts
- Categorizing AQI data into **health-based risk levels** (Good, Moderate, Poor, Very Poor)
- Providing visual cues and simplified terminology to convey **when the air is safe, cautionary, or hazardous**

3. Key Questions on Data

Question 1: Pollutant Frequency Analysis

"Which pollutants show high concentrations most frequently throughout the year?"

This question forms the foundation for understanding Calgary's air pollution profile by identifying the **primary pollutants** that consistently appear in elevated concentrations. The analysis revealed that **particulate matter**, especially **PM2.5** (particles ≤2.5 microns in diameter) and **PM10** (particles ≤10 microns), are the most frequently recorded pollutants at concerning levels. These fine particles originate from various sources such as vehicular emissions, construction dust, industrial activities, residential wood burning, and wildfire smoke.

The high frequency of PM2.5 and PM10 events is alarming from a **public health perspective**. **PM2.5 particles**, in particular, are approximately **30 times smaller than the width of a human hair**, making them capable of penetrating deep into the alveolar regions of the lungs. Once inhaled, these particles can:

- Irritate the respiratory tract
- Trigger or worsen **asthma**, **chronic obstructive pulmonary disease (COPD)**, and **bronchitis**
- Contribute to **heart disease** by inducing systemic inflammation and oxidative stress
- Increase the risk of **hospital admissions** and **mortality** in sensitive groups

Through frequency counts and comparative visualization in Power BI (bar charts), it was determined that PM2.5 events occurred at high levels on a **monthly basis**, particularly in the warmer months, correlating with construction activities and wildfire impacts. PM10, while more prevalent in dry and windy conditions, also showed high concentrations at specific points throughout the year.

In contrast, pollutants like **carbon monoxide (CO)** and **nitrogen dioxide (NO_2)** were less frequently observed at dangerous levels but still showed notable spikes near areas of high traffic density or during temperature inversion events in winter.

This analysis not only identifies **pollutant priorities for mitigation** but also lays the groundwork for **health advisory planning** and regulatory oversight.

Question 2: Temporal Variation Analysis

"How does the Air Quality Health Index (AQHI) change over time in Calgary?"

The **Air Quality Health Index (AQHI)** is a vital composite indicator that translates complex pollutant data into a scale of **low to high health risk**. It incorporates concentrations of key pollutants: **ozone (O_3)**, **nitrogen dioxide (NO_2)**, **sulfur dioxide (SO_2)**, and **fine particulate matter (PM2.5)**. This index allows the public and health professionals to understand the health implications of air quality on a daily basis.

In this case study, **temporal analysis of the AQHI was conducted using monthly trend lines**, highlighting variations throughout the calendar year. The results revealed clear **fluctuations in AQHI levels**, with notable peaks in **late summer and early fall**. These peaks corresponded with known events such as **wildfire activity**, **dry weather**, and **low wind circulation**, which tend to trap pollutants near the surface.

The temporal analysis is especially critical for **vulnerable groups**, including:

- **Elderly individuals (65+)** with reduced pulmonary function
- **Children**, whose developing lungs are more susceptible
- **Pregnant people**, as prenatal exposure can affect fetal development
- Individuals with **asthma, cardiovascular disease, or diabetes**

Periods of elevated AQHI should trigger **public health warnings**, advisories to reduce outdoor activity, and guidance for schools and childcare centers. Moreover, identifying these patterns helps municipal authorities schedule **pollution control operations**, such as reducing vehicular traffic or halting open burning during high-risk days.

In sum, this question explores **how air pollution evolves over time**, enabling **proactive management** rather than reactive mitigation.

Question 3: Seasonal Pattern Investigation

"Are there seasonal patterns in air quality, and what factors contribute to these variations?"

Seasonal variability is a critical dimension in understanding urban air pollution. In Calgary, **distinct seasonal trends** in air quality were observed, shaped by an interplay of **meteorological phenomena**, **human activity**, and **natural events**. This section of the analysis aimed to uncover **how and why air quality changes across seasons** and what can be inferred from these patterns.

Summer Trends: During the **summer months (June to September)**, AQHI values often increase due to:

- **Wildfire smoke** carried from nearby provinces or even the U.S.
- **Elevated ground-level ozone (O_3)** levels, formed from photochemical reactions involving sunlight, NO_x, and VOCs (volatile organic compounds)
- Increased **construction activity** and traffic congestion
- **Temperature inversions** trapping pollutants close to the surface during calm, hot days

In 2024, **September** was identified as the month with the **highest AQI readings**, likely a result of **combined effects** from prolonged dry spells, wildfire smoke drift, and stagnant air conditions.

Winter Trends: Conversely, **winter months (December to February)** displayed variable AQHI behavior, with **some improvement** in pollutant levels like ozone, but occasional spikes in:

- **Nitrogen dioxide (NO_2)** from vehicle emissions, especially during rush hours
- **Carbon monoxide (CO)** from home heating, fireplaces, and idling vehicles
- **Temperature inversions**, where colder air near the surface gets trapped under a layer of warmer air, preventing pollutant dispersion

Winter patterns tend to reflect **localized pollution** from combustion-related sources, exacerbated by Calgary's geography and weather.

Spring and Fall Transitions: In **spring and autumn**, AQI levels showed transitional behavior, but these periods also recorded **intermittent spikes** due to dust storms, agricultural burning, and sudden shifts in weather patterns.

4. Metrics and Key Performance Indicators (KPIs)

To assess and interpret Calgary’s air quality performance in 2024, the study defined a robust framework of **metrics and key performance indicators (KPIs)**. These KPIs were designed to provide both a **quantitative assessment** of pollution levels and a **qualitative classification** of public health impact. Together, these indicators enable stakeholders to make informed decisions about environmental health risks, identify critical trends, and evaluate the success of pollution mitigation strategies.

1. Primary KPIs: Core Air Quality Indicators

Three fundamental KPIs were established to serve as **baseline and extreme-case metrics** for Calgary's air quality throughout the year:

Average AQI (2024)

- Represents the **mean daily Air Quality Index** across all 365 days in 2024.
- Offers a **generalized understanding** of typical ambient air quality conditions.
- Useful for **benchmarking Calgary** against other cities or historical data.
- Helps in assessing whether Calgary’s air quality remains within acceptable levels or requires intervention.

Maximum AQI (2024)

- Identifies the **highest single-day AQI value** during the year.
- Highlights **peak pollution events**, often linked to wildfires, industrial emissions, or stagnant weather patterns.
- These events are of **critical public health importance**, often requiring emergency responses or air quality advisories.
- Used to assess whether mitigation policies are effective during extreme conditions.

Minimum AQI (2024)

- Captures the **lowest AQI value** recorded in 2024.
- Reflects the **cleanest air quality conditions**, often during post-rainfall days, low traffic periods, or strong ventilation.
- Serves as a **performance target** for environmental goals and public awareness campaigns promoting cleaner air.

2. Air Quality Classification Metrics: Health Risk Categories

In alignment with Canadian and international air quality standards, daily AQI values were categorized into **qualitative classes** to better communicate health risks to the public and policymakers. These categories reflect not just pollution levels but their **potential physiological effects** on different population groups:

Category	AQI Range	Health Impact Summary
Good	0–50	Air quality is satisfactory. Air pollution poses little or no risk.
Moderate	51–100	Acceptable air quality. Sensitive individuals may experience minor symptoms.
Unhealthy for Sensitive Groups	101–150	Sensitive groups (e.g., people with asthma, elderly, children) may be affected.
Unhealthy	151–200	Everyone may begin to experience health effects; sensitive groups may experience more serious effects.
Very Unhealthy	201–300	Health alert: The risk of health effects is increased for everyone.
Hazardous	301–500	Health warning of emergency conditions. Everyone is more likely to be affected.

3. Pollutant-Specific Indicators: Targeted Health Risk Metrics

In addition to aggregate AQI values, the study placed significant emphasis on **tracking individual pollutants**, particularly **PM2.5** and **PM10**, due to their **severe impact on human health**, especially respiratory and cardiovascular systems.

PM2.5 (Particulate Matter ≤2.5 μm)

- These ultrafine particles can **bypass nasal and throat filters**, reaching deep into the lungs and even entering the bloodstream.
- Associated with **immediate symptoms** such as coughing, wheezing, shortness of breath, and long-term diseases like asthma, bronchitis, and heart attacks.
- Frequently recorded in high concentrations during **wildfire events**, **traffic congestion**, and **industrial emissions**.

PM10 (Particulate Matter ≤10 µm)

- Slightly larger particles that can still enter the **upper respiratory tract** and cause irritation.
- Elevated during **dry, windy conditions**, **construction activity**, or **road dust resuspension**.
- Linked to throat irritation, sneezing, and exacerbation of existing lung conditions.

Other Pollutants Tracked:

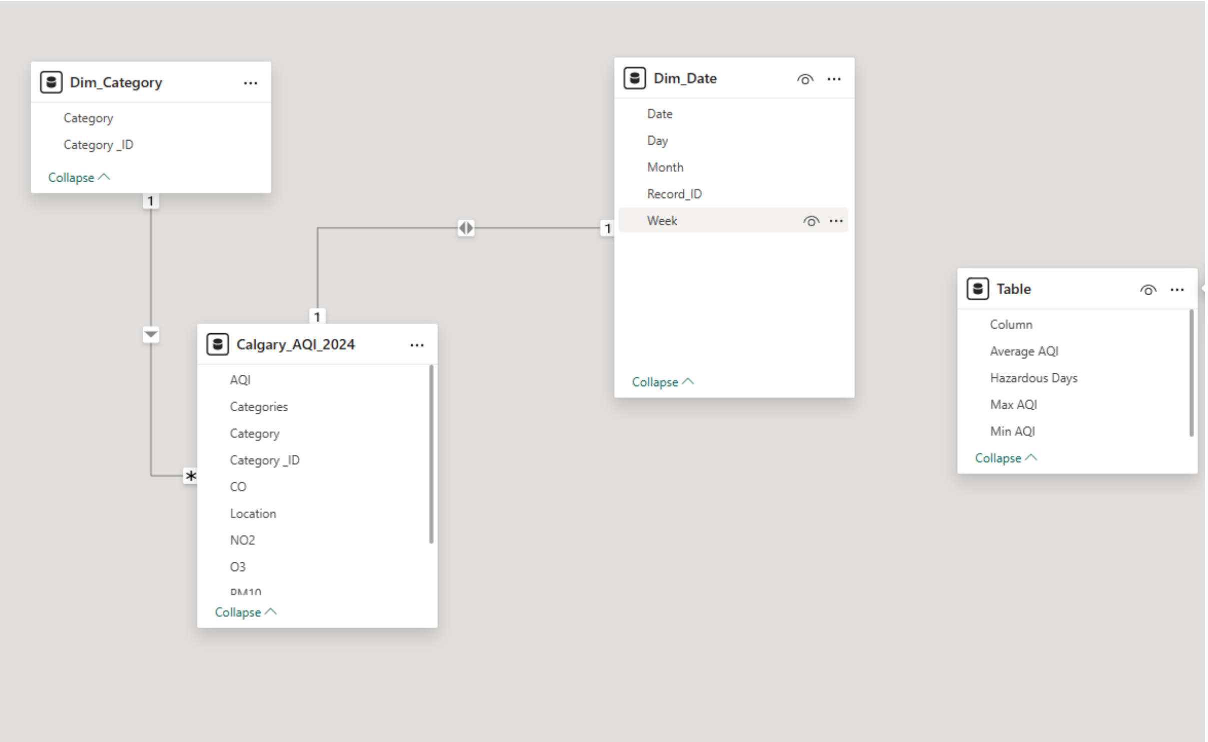
- **NO_2 (Nitrogen Dioxide):** Originates from vehicles and combustion sources; can trigger respiratory inflammation.
- **CO (Carbon Monoxide):** Binds with hemoglobin, reducing oxygen delivery; particularly dangerous for individuals with heart disease.
- **O_3 (Ozone):** A secondary pollutant formed through sunlight-driven reactions; causes chest tightness and lung irritation.

Each pollutant was tracked **individually over time**, using Power BI bar graphs and trend lines to understand **when and where** their concentrations surpassed **health-based thresholds**. This provided a **granular view** of pollutant dynamics and helped isolate sources of specific health risks.

KPI Category	**Purpose**
Primary KPIs	Track overall performance and extreme values
Classification Metrics	Assess public health risk and guide communication
Pollutant-Specific Indicators	Pinpoint health hazard sources and enable targeted action

5. Data Model Architecture

To conduct an effective and scalable analysis of Calgary's air quality data for 2024, a structured **data model** was developed in **Power BI**, using a **star schema design**. This architecture separates data into **fact and dimension tables**, enabling optimized filtering, performance, and dashboard responsiveness. Below is a detailed breakdown of the model and its relationships:



1. Fact Table: Calgary_AQI_2024

This is the **central table** of the model and contains the **daily observational data** related to Calgary's air quality. It includes both AQI values and the pollutant concentrations measured throughout 2024.

Key Columns:

- AQI: The daily Air Quality Index value
- Category: The descriptive classification of AQI (e.g., Good, Moderate)
- Category_ID: The numeric identifier used to link to the category dimension
- PM2.5, PM10: Particulate matter concentrations
- NO2, CO, O3: Gas pollutant concentrations
- Location: Geographic identifier (in this case, Calgary)

Function:

This table serves as the **"fact table"** and is at the center of the model, linking to relevant dimensions to allow filtering, aggregation, and categorization across various attributes like **date** and **category**.

2. Dimension Table: Dim_Category: This dimension table maps numerical category IDs to human-readable category names such as "Good", "Moderate", "Poor", etc.

Key Columns:

- Category_ID: The primary key used to link to the Calgary_AQI_2024 fact table
- Category: The textual label for air quality categories

Purpose:

- Simplifies visual filtering and labeling by replacing numeric codes with **intuitive names**
- Enables grouping and analysis by **air quality levels**, helping in public health risk assessments

Relationship: A **one-to-many** relationship is established between Dim_Category and Calgary_AQI_2024 via Category_ID.

3. Dimension Table: Dim_Date

This dimension supports **time-based slicing and filtering** of AQI data. It enables the analysis of trends by **day**, **month**, and **week**.

Key Columns:

- Date: The full date value (acts as the primary key)
- Day, Month, Week: Decomposed time units for detailed trend analysis
- Record_ID: May serve as a bridge or for redundancy (can be further evaluated)

Purpose:

- Facilitates **monthly**, **weekly**, and **daily trend analysis**
- Enables Power BI time intelligence features like YTD calculations, seasonal comparisons, etc.

Relationship:

- A **one-to-many** relationship is created between Dim_Date and Calgary_AQI_2024 via the Date field.

4. Summary Table: Table (KPI Table)

This appears to be a **disconnected summary table** containing **pre-computed KPI metrics**. It is not directly linked to other tables in the data model but is used for **display purposes** on dashboards and reports.

Key Columns:

- Average AQI: Mean AQI value for 2024
- Max AQI: Highest daily AQI recorded
- Min AQI: Best-case scenario with the lowest AQI value
- Hazardous Days: Count or percentage of days that exceeded a specific AQI threshold

Purpose:

- Provides **standalone metrics** for easy reference and dashboard display
- Useful for creating **card visuals** and highlighting key insights at a glance

Overall Schema Structure and Design Benefits:

Feature	**Benefit**
Star Schema Design	Ensures high performance, modular scalability, and simplified queries
Separation of Concerns	Dimensions handle categories and time logic; fact table handles raw data
Clear Relationships	Enables filtering and drill-down through intuitive connections
Disconnected KPI Table	Supports summary visuals without affecting filtering logic

6. Results

Overall Air Quality Assessment

The 2024 analysis reveals concerning trends in Calgary's air quality performance1. The majority of days throughout the year were classified as either "poor" or "moderate," with only 13% of days achieving "good" air quality status1. This finding indicates persistent challenges in maintaining healthy air quality standards for Calgary residents.

Pollutant Analysis

PM2.5 and PM10 particulate matter emerged as the primary contributors to Calgary's air pollution burden1. These findings align with broader research on urban air quality, where particulate matter consistently represents the most significant health threat. The prevalence of elevated particulate matter levels suggests multiple contributing sources, including vehicular emissions, industrial activities, and seasonal factors like wildfire smoke.

Seasonal Patterns

September demonstrated the highest AQI values of the year, representing a significant seasonal spike in air pollution1. This pattern may reflect several contributing factors:

- **Wildfire Activity**: Late summer wildfires can significantly impact air quality across Alberta, as evidenced by recent events affecting Calgary
- **Meteorological Conditions**: Seasonal wind patterns and atmospheric conditions can concentrate pollutants
- **Industrial Activity**: Potential increases in certain industrial processes during specific periods

Geographic and Temporal Variations

Research indicates that Calgary's air quality varies significantly across different areas of the city and changes with seasonal patterns. Summer pollution tends to concentrate in the downtown core and along major transportation corridors like Crowchild Trail and Deerfoot Trail, while winter patterns show more diffused pollution settling over the city's southern areas

7. Conclusions

Primary Findings

The comprehensive analysis of Calgary's 2024 air quality data reveals several critical insights:

1. **Persistent Air Quality Challenges**: With only 13% of days achieving "good" air quality, Calgary faces ongoing environmental health challenges that require sustained attention1.
2. **Particulate Matter Dominance**: PM2.5 and PM10 particles represent the primary threat to air quality, necessitating targeted interventions addressing their sources1.
3. **Seasonal Vulnerability**: The September spike in air quality issues highlights the need for seasonal preparedness, particularly regarding wildfire impacts and other environmental factors1.

Health Implications

The prevalence of poor and moderate air quality days has significant public health implications Prolonged exposure to elevated PM2.5 and PM10 levels can exacerbate respiratory conditions, increase cardiovascular risks, and particularly impact vulnerable populations including children, elderly individuals, and those with pre-existing health conditions.

Policy and Management Recommendations

The findings suggest several areas for focused intervention:

1. **Source Control**: Addressing the primary sources of particulate matter through enhanced emissions standards and industrial monitoring
2. **Public Health Preparedness**: Developing robust early warning systems for high pollution days, particularly during wildfire season
3. **Urban Planning**: Considering air quality impacts in transportation and development planning to reduce pollution hotspots

Future Research Directions

This analysis establishes a foundation for ongoing air quality monitoring and research. Future investigations should explore:

- **Source Attribution**: Detailed analysis of specific pollution sources contributing to elevated PM levels
- **Health Impact Assessment**: Quantitative evaluation of health outcomes associated with observed air quality patterns
- **Intervention Effectiveness**: Assessment of policy measures designed to improve air quality outcomes