Assignment 1: Air Quality Data Analysis

The data for this analysis was collected from the Vancouver International Airport #2 - Air Monitoring Station located on 49.193333, -123.175659. This station monitors various pollutants, including CO, O3, NO2, and PM2.5, providing valuable insights into the air quality at and around the Vancouver International Airport.

Table of Contents

Cleaned Dataset

The `cleaned-dataset` is shown in figure 1

	Date	Time	TEMP_MEAN_°C	CO_ppm	NO2_ppb	O3_ppb	PM2.5_ug/m3	SO2_ppb	PM10_ug/m3	NO_ppb	WSPD_VECT_m/s	WDIR_VECT_Deg.
0	12/31/2022	24:00 AM	4.4	0.23	16.7	6.8	5.2	0.2	5.9	2.7	0.76	284.3
1	1/1/2023	1:00 AM	4.2	0.29	22	3.7	8.7	0.3	9	7.3	0.46	184
2	1/1/2023	2:00 AM	4	0.31	23.5	1.9	21.4	0.4	12.1	17.5	0.84	321.6
3	1/1/2023	3:00 AM	4.2	0.29	19.5	2.3	21.6	0.3	15.1	4.8	0.81	4.6
4	1/1/2023	4:00 AM	4.4	0.29	19	1.5	15.2	0.3	12.1	7.3	0.69	234.7

Figure 1: Cleaned dataset

Understanding Data Structure

Figure 2 shows the data types of each column, which are currently string however we have datetime and floating point values. after convertion our new data structure is shown below.

```
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Date            8220 non-null   object
 1   Time            8220 non-null   object
 2   TEMP_MEAN_°C    8220 non-null   object
 3   CO_ppm          8220 non-null   object
 4   NO2_ppb         8220 non-null   object
 5   O3_ppb          8220 non-null   object
 6   PM2.5_ug/m3     8220 non-null   object
 7   SO2_ppb         8220 non-null   object
 8   PM10_ug/m3      8220 non-null   object
 9   NO_ppb          8220 non-null   object
 10  WSPD_VECT_m/s   8220 non-null   object
 11  WDIR_VECT_Deg.  8220 non-null   object
```

```
#   Column           Non-Null Count  Dtype
---  ------           --------------  -----
0   Date             8220 non-null   datetime64[ns]
1   Time             8220 non-null   object
2   TEMP_MEAN_°C     8220 non-null   float64
3   CO_ppm           8220 non-null   float64
4   NO2_ppb          8220 non-null   float64
5   O3_ppb           8220 non-null   float64
6   PM2.5_ug/m3      8220 non-null   float64
7   SO2_ppb          8220 non-null   float64
8   PM10_ug/m3       8220 non-null   float64
9   NO_ppb           8220 non-null   float64
10  WSPD_VECT_m/s    8220 non-null   float64
11  WDIR_VECT_Deg.   8220 non-null   float64
```

Finding negative values

	Date	Time	TEMP_MEAN_°C	CO_ppm	NO2_ppb	O3_ppb	PM2.5_ug/m3	SO2_ppb	PM10_ug/m3	NO_ppb	WSPD_VECT_m/s	WDIR_VECT_Deg.
674	2023-01-29	02:00	-0.3	0.15	12.5	23.4	2.0	0.4	7.2	0.1	2.00	77.9
675	2023-01-29	03:00	-0.5	0.18	14.7	20.6	1.5	0.3	7.3	0.0	1.14	94.2
676	2023-01-29	04:00	-1.0	0.21	17.2	17.8	1.9	0.3	7.6	0.0	1.52	88.4
677	2023-01-29	05:00	-1.3	0.20	17.1	15.6	1.7	0.3	7.8	0.3	1.26	49.8
678	2023-01-29	06:00	-1.7	0.23	21.5	12.0	1.9	0.3	9.6	0.3	1.00	65.7

Figure 2: Dataset showing negative reading

Practically, negative readings may arise from sensor data due to several issues such as calibration errors, power failures, and sensor limitations. We can address these issues by understanding the behavior of our measured parameters. For instance, it is unlikely for readings of CO, CO2, or similar parameters to be negative. Therefore, we filter out any negative readings. About 255 rows have negatives, this represent a small proportion of our dataset approximately 1.1%, dropping them is unlikely to significantly impact the overall analysis.

Statistical Analysis

Descriptive Statistics

The summary statistics of the dataset show 7965 records of air quality measurements. Key observations include mean CO levels at 0.204 ppm and mean NO levels at 7.09 ppb. Notably, NO levels vary widely (std: 15.68 ppb), with maximum readings reaching 201.2 ppb. PM2.5 averages 5.12 µg/m³.

	Date	TEMP_MEAN_°C	CO_ppm	NO2_ppb	O3_ppb	PM2.5_ug/m3	SO2_ppb	PM10_ug/m3	NO_ppb	WSPD_VECT_m/s	WDIR_VECT_Deg.
count	7965	7965.000000	7965.000000	7965.000000	7965.000000	7965.00000	7965.000000	7965.000000	7965.000000	7965.000000	7965.000000
mean	2023-07-06 01:15:23.389830656	11.167458	0.203694	11.603478	19.161908	5.12791	0.439573	11.343528	6.945888	2.624203	162.572982
min	2023-01-01 00:00:00	0.000000	0.020000	0.600000	0.300000	0.00000	0.100000	0.000000	0.000000	0.040000	0.000000
25%	2023-04-11 00:00:00	6.300000	0.140000	5.200000	8.900000	2.40000	0.200000	7.200000	0.300000	1.530000	90.500000
50%	2023-07-05 00:00:00	10.600000	0.170000	9.600000	19.500000	4.10000	0.300000	9.900000	1.500000	2.440000	122.600000
75%	2023-09-30 00:00:00	15.800000	0.230000	16.800000	28.600000	6.40000	0.500000	13.500000	5.900000	3.390000	261.800000
max	2023-12-30 00:00:00	27.900000	1.180000	53.600000	55.000000	87.40000	12.800000	197.400000	201.200000	12.580000	359.900000
std	NaN	5.860527	0.105623	7.963532	12.232996	4.37524	0.355961	7.707613	15.384664	1.544653	93.024409

Figure 3: Descriptive Statistics

Time Series Analysis

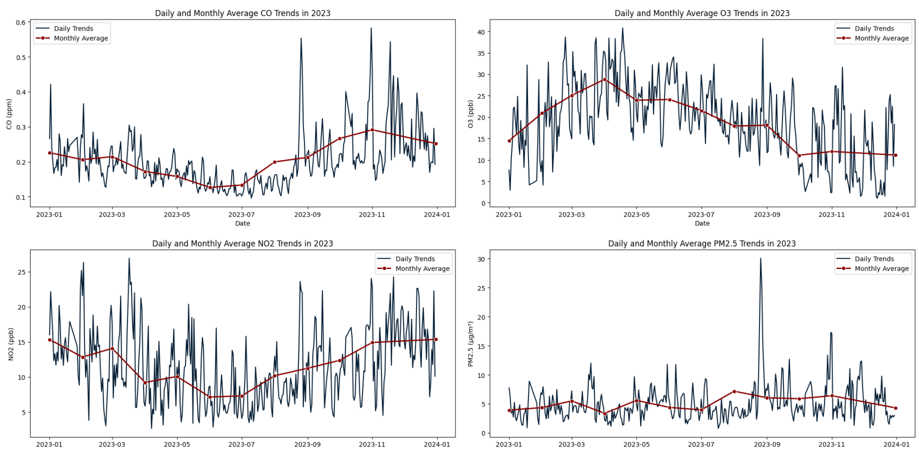


Figure 4: Time Series Analysis

Daily/Monthly Trend Analysis

Daily CO Trends

- **Description**: This chart displays the concentration of Carbon Monoxide (CO) in parts per million (ppm) over the months.
- **Observations**:
  - The CO levels starts at a high of **0.23 ppm** in **January** and show a general decreasing trend until **June**.
  - After June, CO levels begin to increase, peaking in **November** of around **0.29 ppm** before slightly dropping towards **December**.

Daily 03 Trends

- **Description**: This chart displays the concentration of Ozone (O3) in parts per billion (ppb) over the months.
- **Observations**:
  - Ozone levels start low of **13.98 ppb** in **January** and sharply rose, reaching a peak of **28.34 ppb** in **April**.

- Post-April, there is a gradual decline in O3 levels, hitting a significant low in **October 10.88 ppb** before remaining farily stable towards the end of the year.

Daily NO2 Trends

- **Description**: This chart shows the concentration of Nitrogen Dioxide (NO2) in parts per billion (ppb) over the months.
- **Observations**:
  - NO2 levels are high at the start of the year of around 1**5.85 ppb** in **January**, with fluctuations observed until **April**.
  - A downward trend is seen from **April**, reaching the lowest in **June 7.12 ppb**.
  - After **June**, NO2 levels rises steadily, ending the year at a higher level.

Daily PM25 Trends

- **Description**: This chart depicts the concentration of Particulate Matter (PM2.5) in micrograms per cubic meter (µg/m^3) over the months.
- **Observations**:
  - PM2.5 levels show significant variation throughout the year. Starting at **3.96 µg/m^3** in **January**, the levels peak in **August 7.17 µg/m^3**, indicating higher particulate matter concentration.
  - The lowest concentration is observed in April **(3.51 µg/m^3)**

Summary

The charts illustrate the seasonal and Daily fluctuations in air quality indicators. The patterns reveal periods of elevated and reduced concentrations for each pollutant, indicating potential seasonal influences or diverse emission sources. For example, carbon monoxide (CO) and nitrogen dioxide (NO2) levels typically rise during colder months, possibly due to increased heating activities couple with airport located near by, while ozone (O3) peaks in spring, likely because of photochemical reactions facilitated by greater sunlight exposure.

Correlation Analysis

The correlation matrix reveals strong positive correlations between CO, NO, and PM2.5, indicating shared sources or similar emission patterns. In contrast, CO and O3 show a negative correlation, indicating different formation mechanisms. Temperature shows a negative correlation with CO, NO, and PM2.5, and a positive correlation with O3.

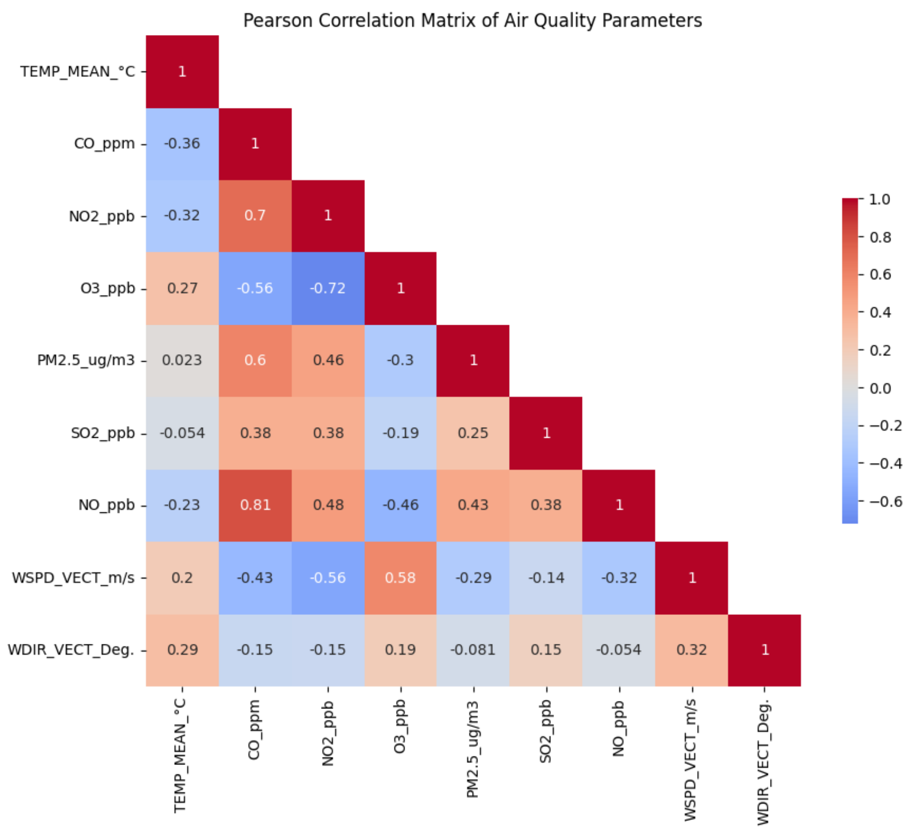


Figure 5: Correlation

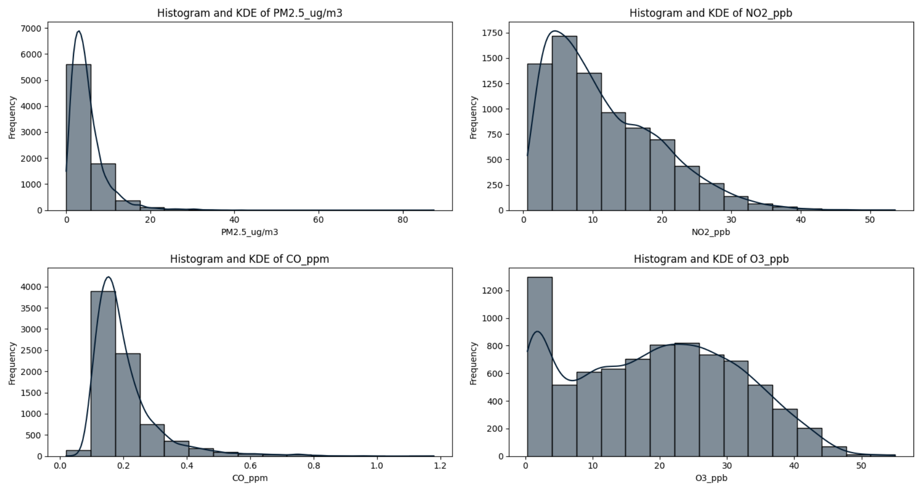


Figure 6: Distribution

Distribution Analysis

- CO (Carbon Monoxide): Concentrations are skewed towards lower values, with most readings below 0.3 ppm.
- NO2 (Nitrogen Dioxide): Shows a distribution with a peak around 8 ppb, indicating common levels around this value.
- O3 (Ozone): Shows a relatively spread distribution, with multiple peaks suggesting variability in readings.
- PM2.5: Concentrations are mostly below 20 µg/m³, with a right-skewed distribution indicating occasional high readings.

Air Quality Health Index

	Date	TEMP_MEAN_°C	CO_ppm	NO2_ppb	O3_ppb	PM2.5_ug/m3	SO2_ppb	PM10_ug/m3	NO_ppb	WSPD_VECT_m/s	WDIR_VECT_Deg.	color	aqhi_general
0	2023-01-01	4.708000	0.266000	16.012000	7.652000	7.736000	0.384000	7.460000	5.652000	1.280800	125.592000	#00BFFF	0.888503
1	2023-01-02	2.239130	0.420870	22.156522	2.917391	6.304348	0.452174	8.300000	29.095652	1.058261	127.417391	#00BFFF	0.643317
2	2023-01-03	3.320833	0.242083	19.245833	10.279167	3.383333	0.295833	6.066667	4.941667	2.453750	80.895833	#00BFFF	0.870082
3	2023-01-04	5.181818	0.200909	16.259091	13.768182	4.022727	0.322727	5.318182	4.586364	2.591364	90.272727	#1b8fb5	1.048946
4	2023-01-05	8.479167	0.166667	12.329167	22.020833	2.712500	0.270833	7.620833	1.000000	3.134583	104.466667	#1b8fb5	1.383552

Figure 7: AQHI

The above data illustrates the daily trends in AQHI (Air Quality Health Index). A comparison with regulatory standards will be conducted below, referencing guidelines from Environment and Climate Change Canada.

Regulatory Comparison

Air Quality Health Index (AQHI) and Regulatory Compliance in British Columbia The Air Quality Health Index (AQHI) is designed to protect public health by offering real-time air quality data and health recommendations. In Canada, Health Canada and Environment and Climate Change Canada work together to oversee the AQHI, monitoring air quality and educating the public about potential health hazards. These regulations are crucial for shielding vulnerable populations from the harmful impacts of air pollution.

The AQHI uses a color-coded system to communicate the level of health risk associated with air quality. According to Environment and Climate Change Canada, the AQHI ranges are as follows :

- Low Risk (1-3): Colors ranging from light blue to dark blue.
- Moderate Risk (4-6): Yellow to gold and orange.
- High Risk (7-10): Shades of red.
- Very High Risk (10+): Dark red.

In a study published in PLOS ONE, Dunbar et al. discuss the significance of regulatory compliance in health services, emphasizing the mandatory adherence to standards set by regulatory bodies to ensure safety and quality. This compliance often involves inspections and sanctions for non-compliance. Additionally, a systematic review in Health Policy and Planning highlights that effective regulatory frameworks are critical for improving system performance and ensuring compliance through structured governance and regular evaluations . These frameworks play a vital role in maintaining air quality standards and protecting public health.

In British Columbia, regulatory limits for key pollutants are set as follows : - PM2.5: The limit is 25 $\mu g/m^3$ over a 24-hour period, with no 1-hour objective set by the provincial government. - NO2: The goal is set at 60 parts per billion (ppb) hourly. - O3 (Ozone): The limit is 62 ppb over an 8-hour period (Environment BC).

These limits help regulate and maintain air quality standards, thereby mitigating the adverse health effects of air pollution on the population.

Comparison of pollutants concentration to BC objective

The chart provides an analysis of air quality trends in relation to British Columbia's regulatory objectives for PM2.5, O3, and NO2. The top left panel indicates daily PM2.5 levels, which generally remain below the BC objective of 25 $\mu g/m^3$, with occasional spikes exceeding this limit. The top right panel illustrates the 8-hour rolling average of O3 levels, which are below the BC objective of 62 ppb. The bottom panel displays hourly NO2 trends, typically staying below the 60 ppb objective. These visualizations help in identifying periods when pollutant levels exceed regulatory limits, supporting air quality management and public health protection. In future sections, we will explore insight into the occasional spike in PM2.5.

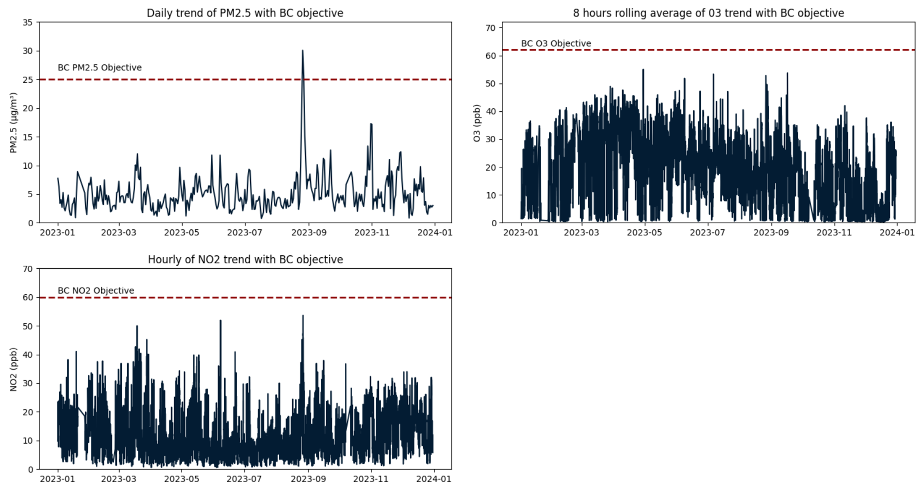


Figure 8: Comparison of pollutants against objective

AQHI Comparison to Environment and Climate Change Canada objective

The Air Quality Health Index (AQHI) chart displays the variation in air quality over time from January 2023 to December 2023. Each point represents the AQHI value for a specific day, with colors indicating the risk level associated with air quality. Most points cluster around values of 1 to 2, indicating generally low health risk around Vancouver International Airport #2 adhering to the Environment and Climate Change Canada objective for AQHI.

Major Emission Sources

Major Emission Sources and Wind Direction in Richmond, BC

Richmond, BC, is a coastal city with significant industrial activity, including the Richmond Plywood Factory and the Richmond Cement Plant. These facilities release pollutants such as particulate matter (PM2.5), nitrogen oxides (NOx), and carbon monoxide (CO), impacting local air quality.

Monthly data from these sources show the following emissions:

Richmond Plywood Factory: - PM2.5 emissions range from 1.642900 to 2.229650 tons per month. - NOx emissions range from 5.03370 to 6.83145 tons per month. - CO emissions range from 10.159800 to 13.788300 tons per month.

Richmond Cement Plant: - PM2.5 emissions are consistently around 0.871318 to 0.872364 tons per month. - NOx emissions are consistently around 93.04610 to 93.15780 tons per month. - CO emissions are consistently around 45.195248 to 45.249504 tons per month.

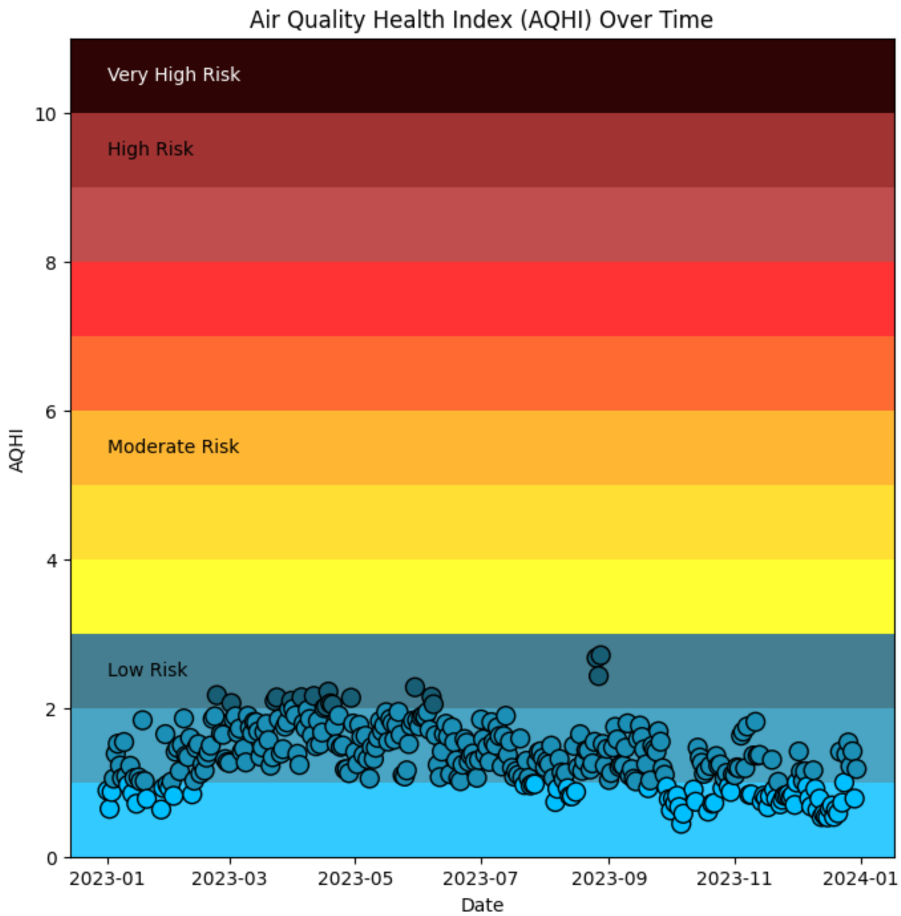


Figure 9: Comparison of AQHI with

Wind patterns are crucial in dispersing these pollutants. Richmond generally experiences moderate winds, with prevailing directions changing seasonally. During summer, winds predominantly blow from the west and northwest, bringing cooler marine air and dispersing pollutants away from the city. In winter, winds often come from the east and southeast, associated with frontal systems from the Pacific Ocean.

Understanding these wind patterns is essential for predicting air quality impacts. The Vancouver monitoring station, strategically located at the coast and downwind from major industries during winter months, effectively monitors air quality and assesses the impact of industrial emissions on public health and the environment.

Emission from mentioned sources

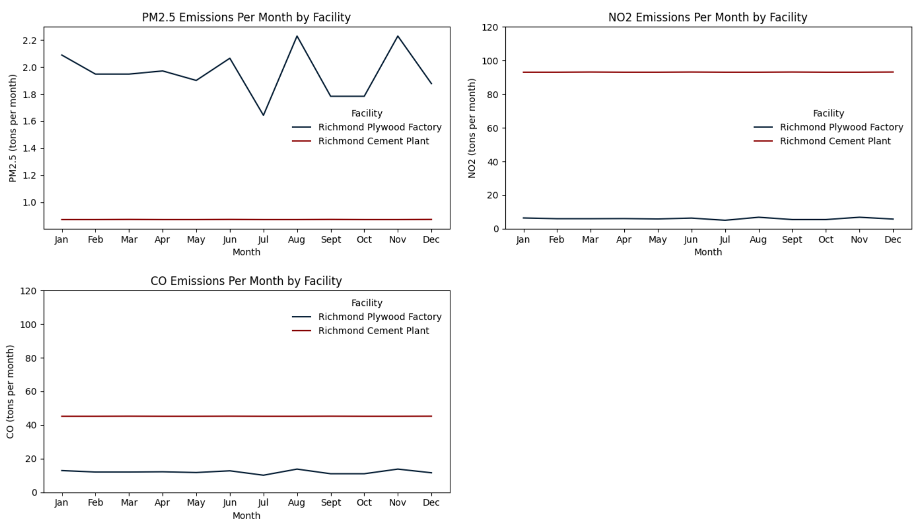
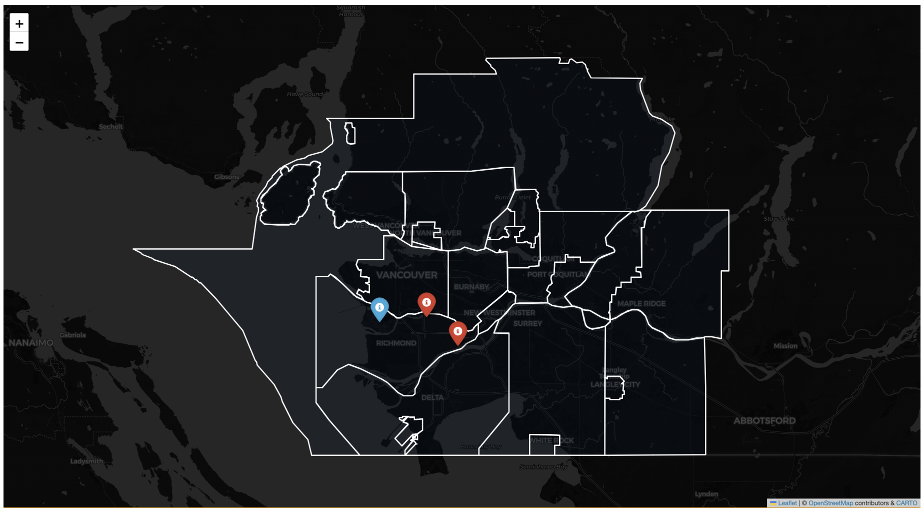


Figure 10: Higlighted emission sources

The chart displays emissions per month for Richmond Plywood Factory and Richmond Cement Plant. The Richmond Plywood Factory shows fluctuating PM2.5,CO, and NO2 emissions, with notable peaks in August and November. These peaks correspond to higher concentrations measured at the monitoring station, likely due to winter southeast winds bringing pollutants inland and exacerbating air quality issues. In contrast, the Richmond Cement Plant maintains consistent emission throughout the year.

Further examination of the pollutant concentation data from station, particularly on the hour with the high pollutant concentration levels, which is in November, a winter month, reveals a wind direction of 86 - 98.5 degrees, which corresponds to an east-southeast wind. This wind direction supports the hypothesis that pollutants are being transported inland, impacting air quality in the monitored area.

Spatial context



The map shows the facilities' geographic locations, with the monitoring station indicated by a blue marker for spatial context.

Summary/Conclusion

This analysis evaluates air quality data from the Vancouver International Airport #2 monitoring station, focusing on pollutants such as CO, NO2, O3, and PM2.5. Thorough data-cleaning procedures were applied to ensure accuracy, and statistical analysis revealed significant trends and correlations among the pollutants. For instance, O3 is positively correlated with temperature, with levels peaking during spring, attributed to photochemical reactions facilitated by greater sunlight exposure. CO, NO2, and PM2.5 show strong correlations with each other. Daily patterns highlighted seasonal variations, with CO and NO2 levels peaking during colder months, likely due to increased heating activities and airport operations, coupled with east-southeast winds blowing pollutants into the city.

Comparing the data against the standards set by Province of British Columbia revealed that pollutant levels generally met acceptable limits. However, occasional spikes in PM2.5 exceeded regulatory thresholds. Spatial analysis identified emission sources with high emission rates during these periods, highlighting the impact of wind patterns on pollutant dispersion. This comprehensive analysis underscores the importance of ongoing air quality monitoring and regulatory compliance to protect public health and effectively manage pollution.