Air Quality Index (AQI) Report

Data Source: West Bengal Pollution Control Board

Station: Bhasa,2nd Campus of Asutosh College

(JUNE_2025)

Introduction

The Air Quality Index (AQI) is a tool developed to represent the quality of air in a simplified and accessible format for the general public. It translates complex data from various air pollutants into a single number or scale, usually ranging from 0 to 500. The higher the AQI value, the greater the level of air pollution and the higher the potential for adverse health effects. The index is divided into categories such as *Good (0–50)*, *Satisfactory (51–100)*, *Moderate (101–200)*, *Poor (201–300)*, *Very Poor (301–400)*, and *Severe (401–500)*, each associated with specific health advisories and color codes.

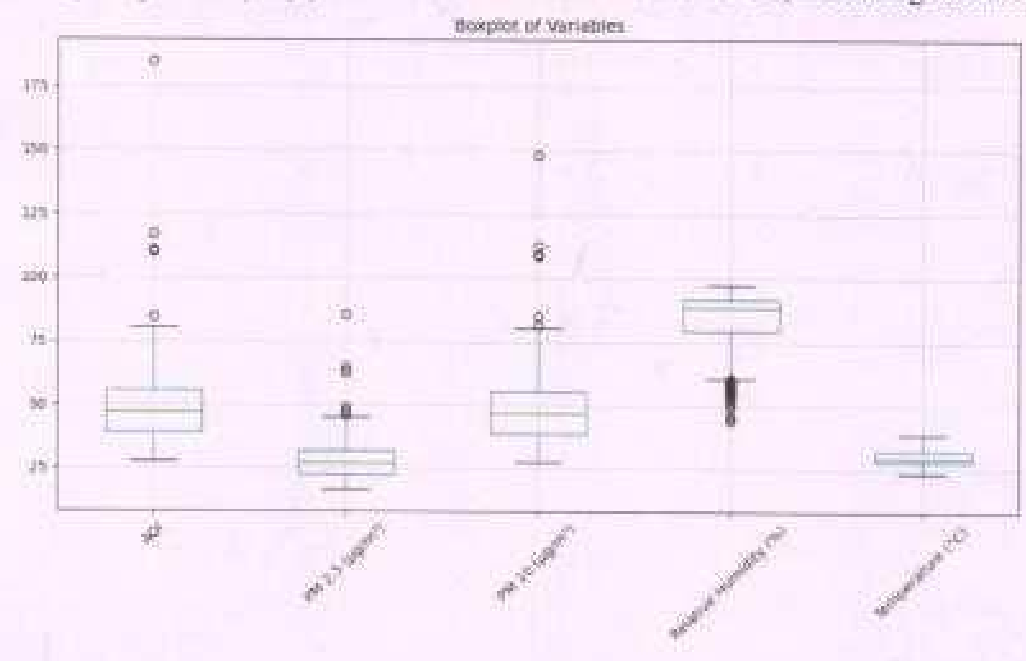
AQI takes into account the concentration of key pollutants including Particulate Matter ($PM_{2.5}$ and PM_{10}), Nitrogen Dioxide (NO_2), Sulphur Dioxide (SO_2), Ozone (O_3), Carbon Monoxide (CO), and Ammonia (NH_3). These pollutants are measured at various monitoring stations and averaged over specific time periods.

The AQI is especially important in urban and industrial areas, where pollution levels tend to fluctuate due to vehicular emissions, construction activities, industrial discharges, and weather conditions. Regular monitoring and dissemination of AQI information help individuals make informed decisions about outdoor activities, especially those with respiratory or cardiovascular issues.

In addition to public health, AQI also serves as a guiding tool for policymakers and environmental agencies to frame air pollution control strategies and implement sustainable development practices. By raising awareness and encouraging behavioral change, the AQI plays a vital role in promoting a cleaner and healthier environment.

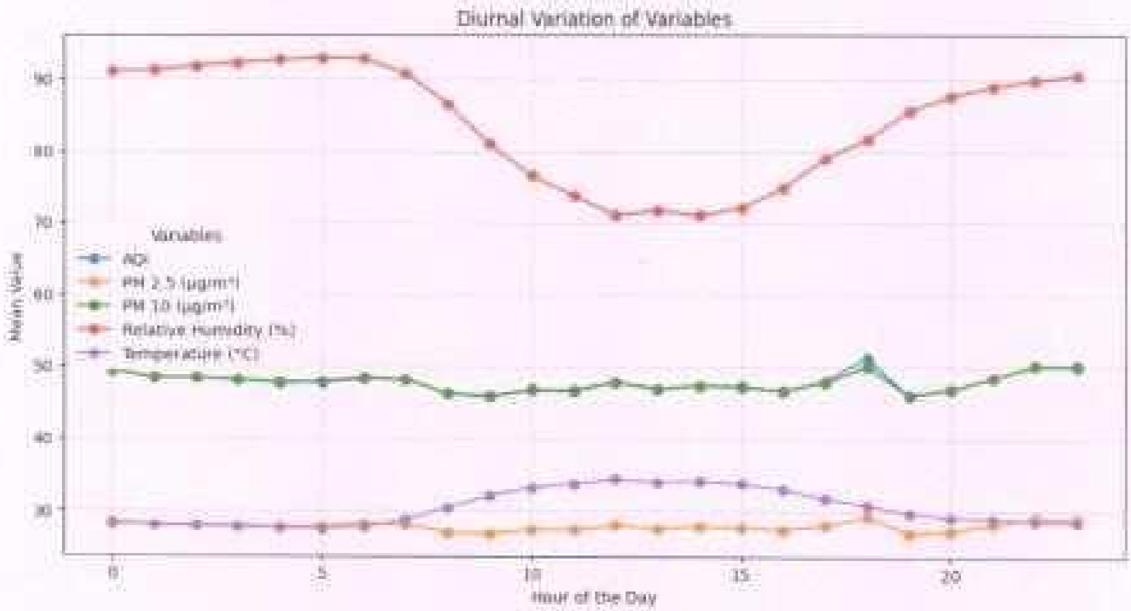
DATA ANALYSIS

The boxplot provides a comparative statistical overview of various atmospheric parameters—AQI, PM 2.5, PM 10, relative humidity, and temperature—measured over a specific period. The Air Quality Index (AQI) shows a median value around 45–50, indicating that the air quality



generally remained in the "moderate" category. However, the presence of several high outliers, some exceeding 180, suggests occasional spikes in pollution levels, potentially due to specific meteorological or anthropogenic activities. PM 2.5 concentrations are relatively lower, with a median around 27 µg/m³, though sporadic high values imply transient episodes of fine particulate pollution. PM 10 displays a wider range and higher variability compared to PM 2.5, with outliers reaching beyond 140 µg/m³, indicating significant influence from dust and coarse particles on certain days.

Relative humidity remains consistently high, with a median close to 90%, suggesting humid atmospheric conditions typical of the monsoon or pre-monsoon season. Although the humidity range is narrow, the presence of low-end outliers below 65% could indicate short dry spells or fluctuations in weather patterns. Temperature values appear stable, with a median around 30°C and a relatively narrow interquartile range, reflecting uniform thermal conditions. A few lower outliers, below 25°C, possibly represent early morning readings or days with rainfall-induced cooling. Overall, the boxplot highlights that while the air quality remained moderate on most days, the influence of particulate matter, especially PM 10, and humidity levels played a crucial role in daily air quality fluctuations.



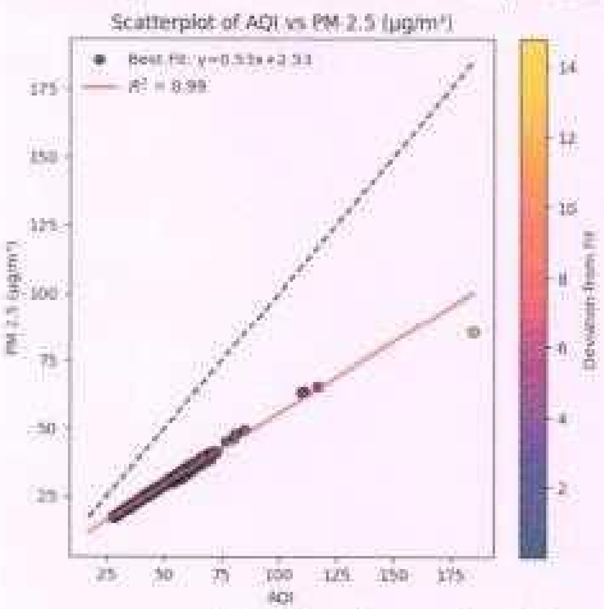
The graph depicts the hourly variation of five key environmental variables—AQI, PM 2.5, PM 10, Relative Humidity (%), and Temperature (°C)—throughout a 24-hour period. The chart provides insights into how atmospheric conditions change over the course of a day, reflecting interactions between pollution levels and meteorological factors.

The diurnal pattern reveals that **relative humidity is highest during early morning hours (around 12 AM to 7 AM)**, peaking above 90%, and steadily declines through the late morning and afternoon, reaching its lowest levels (around 71%) between **12 PM and 3 PM**. This trend then reverses, with humidity rising again in the evening and night due to cooler temperatures and decreased solar radiation. **Temperature** shows a reverse trend: it remains lowest in the early morning (about 28°C), gradually increases, and peaks between **12 PM and 3 PM at nearly 34°C**, before declining again in the evening.

Particulate pollutants such as **PM 2.5 and PM 10** show relatively stable behavior throughout the day, with slight increases during early morning and late evening, likely due to low wind dispersion and increased human activity (e.g., traffic, cooking, etc.). **AQI mirrors the variation of PM concentrations**, staying moderate during the day but experiencing a **sharp rise around 6 PM**, peaking close to 52 around **7 PM**, possibly due to the combination of traffic emissions and lowering atmospheric mixing height during the evening hours. Afterward, the AQI drops again, aligning with pollutant settling and decreased activity.

- Relative **Humidity (%):**
  - Peaks during early morning and late night (>90%).
  - Lowest between 12 PM and 3 PM (~71%).
  - Indicates typical cooling and moisture accumulation pattern.
- Temperature **(°C):**
  - Lowest (~28°C) during 3 AM–6 AM.
  - Peaks (~34°C) during early to mid-afternoon (12 PM–3 PM).
  - Inversely related to humidity.
- PM **2.5 and PM 10 (μg/m³):**
  - Relatively stable, with mild dips during mid-day.
  - Slight increases in early morning and late evening.
  - Suggests local emissions and limited dispersion.
- AQI:
  - Follows a stable trend but shows a **notable spike at 7 PM**, likely from traffic and human activities.
  - Indicates evening deterioration in air quality.

This scatterplot visualizes the relationship between the Air Quality Index (AQI) and PM 2.5 concentrations, with a linear regression model applied for quantitative interpretation.



Strong Positive Correlation:

- The scatterplot shows a **strong linear correlation** between AQI and PM 2.5 concentrations.
- The **best-fit regression line** is given as: PM 2.5=0.53×AQI+2.53\text {PM 2.5} = 0.53 \times \text {AQI} + 2.53PM 2.5=0.53×AQI+2.53
- This indicates that for every unit increase in AQI, PM 2.5 increases by approximately **0.53 μg/m³**.

2. **High Coefficient of Determination (R^2 = 0.99):**
   - An **R^2 value of 0.99** suggests that **99% of the variation in PM 2.5** can be explained by changes in AQI.
   - This implies a **very strong statistical relationship** and indicates that PM 2.5 is a primary contributor to AQI levels at the observed location.
3. **Color Gradient – Deviation from Fit:**
   - The points are colored based on their **deviation from the regression line**.
   - Most data points cluster around the line (darker shades), indicating **low deviation**.

- A few points, particularly one in yellow (top-right), exhibit **larger deviation**, potentially representing an outlier due to local pollution sources or measurement anomalies.

4. **Dashed Line Reference:**
   - The **black dashed line** represents the 1:1 relationship (PM 2.5 = AQI).
   - The fact that the red regression line lies **below** the dashed line shows that **PM 2.5 values are generally lower than the AQI values**, which is expected since AQI incorporates scaling and multiple pollutants.

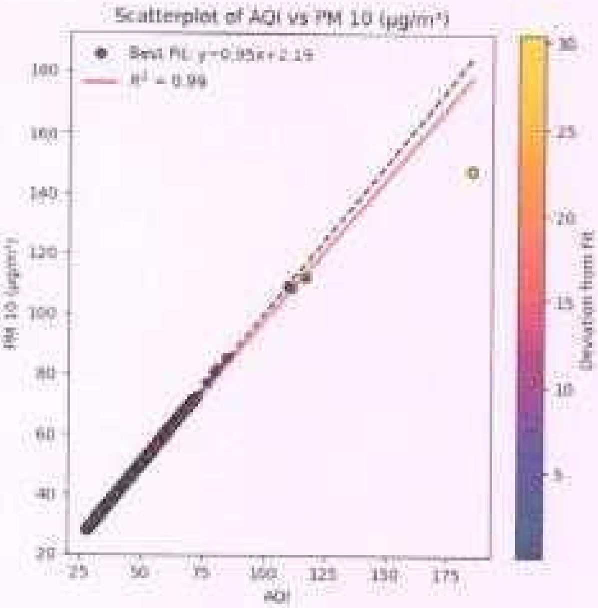
- There is a **very strong linear relationship** between AQI and PM 2.5.
- The **regression model fits the data well**, with $R^2 = 0.99$.
- PM 2.5 is the **dominant pollutant influencing AQI** in this dataset.
- Outliers are minimal, indicating **consistent environmental behaviour**, with only a few extreme deviations.
- The plot confirms that **AQI monitoring can be reliably approximated using PM 2.5 data** in this context.

This scatterplot illustrates the relationship between the **Air Quality Index (AQI)** and **PM 10 concentrations**, supported by a linear regression model to interpret the strength and nature of the correlation.

Detailed Interpretation:

1. **Strong Linear Correlation:**
   - The **best-fit line** has the equation:
     PM 10=0.95×AQI+2.19\text{PM 10} = 0.95 \times \text{AQI} + 2.19PM 10=0.95×AQI+2.19
   - This suggests that **PM 10 increases nearly one-to-one** with AQI, with each unit increase in AQI associated with a **0.95 µg/m³** increase in PM 10 concentration.
   - The slope being close to 1 indicates a **very tight, proportional relationship**.
2. **Very High R^2 Value (0.99):**
   - The **coefficient of determination ($R^2 = 0.99$)** reflects that **99% of the variability in PM 10** can be explained by changes in AQI.
   - This demonstrates that **PM 10 is a dominant factor in AQI formulation** in this dataset.
3. **Deviation Visualization:**
   - Color-coded points represent **deviation from the fitted line**.
   - The majority of data points show **minimal deviation**, indicating excellent fit and data consistency.
   - A noticeable outlier (in yellow, far right) has a **significant positive deviation**, suggesting an unusually high PM 10 level at a given AQI—likely due to a localized pollution event or measurement error.
4. **Comparison with 1:1 Line (Black Dashed Line):**
   - The **dashed line (y = x)** is a reference for equal values of AQI and PM 10.

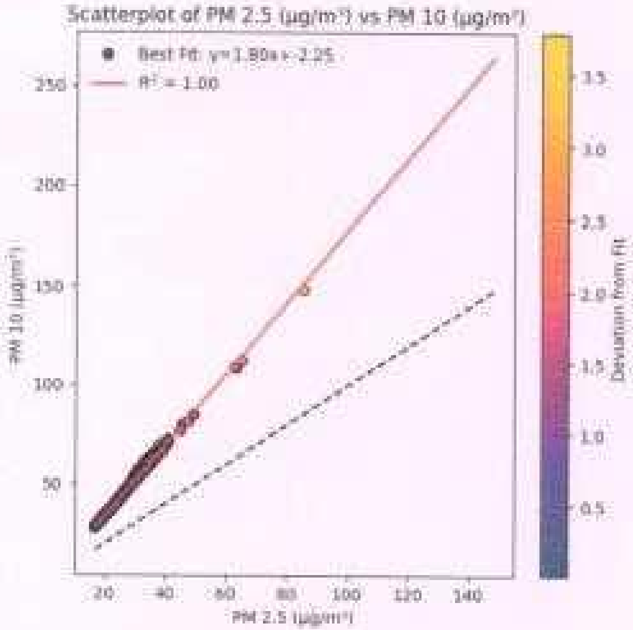
- The regression line falls just below the 1:1 line, reinforcing that **AQI is slightly higher than the PM 10 values** at most points, as AQI is a composite index.



- **Regression equation**: PM 10 = 0.95 × AQI + 2.19
- **R^2 = 0.99**, indicating a **strong, reliable relationship**.
- PM 10 is shown to be **directly proportional** to AQI, contributing significantly to it.
- Most data points lie very close to the fit line, implying **low dispersion and high consistency**.
- One significant **outlier** indicates a rare event or anomaly.
- The linear trend supports using **PM 10 as a predictor** for AQI in the region.

This scatterplot examines the linear relationship between **PM 2.5** and **PM 10** concentrations, two key particulate pollutants in air quality assessment.

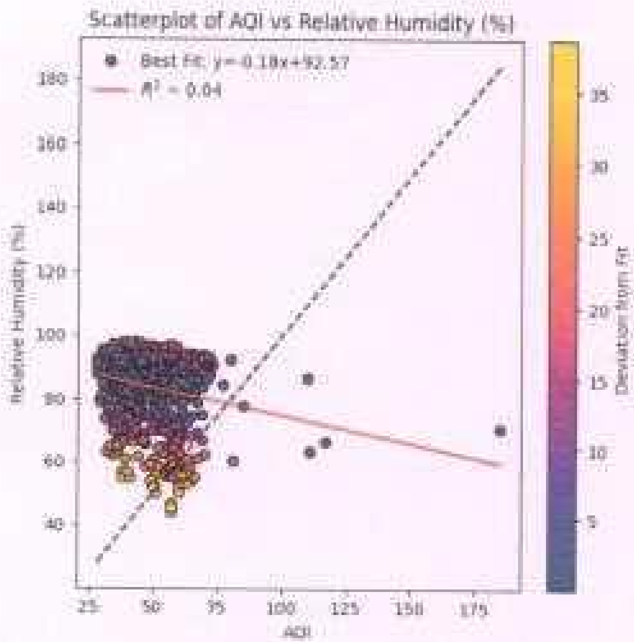
1. **Strong Linear Relationship:**
   - The best-fit line is defined by the equation: PM 10=1.80×PM 2.5−2.25\text{PM 10} = 1.80 \times \text{PM 2.5} - 2.25PM 10=1.80×PM 2.5−2.25
   - This indicates that PM 10 concentrations are **approximately 1.8 times** higher than PM 2.5 concentrations, accounting for the presence of both coarse and fine particles in PM 10.
2. **Perfect Fit (R^2 = 1.00):**
   - The **coefficient of determination (R^2 = 1.00)** shows an **almost perfect linear correlation** between PM 2.5 and PM 10 values.
   - This suggests a very tight coupling, implying that PM 2.5 consistently forms a significant portion of PM 10 in the observed dataset.
3. **Color-Coded Deviation:**
   - The color gradient represents the **deviation from the regression line**.
   - Most points are colored dark (low deviation), indicating **minimal variation** from the linear trend.
   - A few scattered outliers (yellow and orange points) deviate slightly more but are still close to the predicted values.
4. **Comparison with 1:1 Line (Dashed Line):**
   - The **black dashed line (y = x)** shows what the values would be if PM 2.5 and PM 10 were equal.
   - The regression line lies **well above** the 1:1 line, confirming that **PM 10 levels are consistently higher** than PM 2.5, as expected due to PM 10 including both coarse and fine particles.



- **Regression equation:** PM 10 = 1.80 × PM 2.5 − 2.25
- **R^2 = 1.00**, indicating a **nearly perfect correlation**.
- PM 2.5 is a **major component of PM 10**, but PM 10 also includes additional coarse particles.
- Very **low deviation** across data points confirms **consistency** in pollutant composition.
- The scatterplot suggests **predictability** and **strong co-dependence** between these two pollutants.

This scatterplot explores the relationship between **Air Quality Index (AQI)** and **Relative Humidity (%)**, using a linear regression fit and color-coded deviation to assess the nature and strength of correlation.

1. **Weak Negative Correlation:**
   - The regression line equation is:
     Relative Humidity=−0.18×AQI+92.57\text{Relative Humidity} = -0.18 \times \text{AQI} + 92.57Relative Humidity=−0.18×AQI+92.57
   - This indicates a **slight inverse relationship**, meaning that as AQI increases, relative humidity tends to decrease **marginally**.
   - However, the slope is **very shallow**, suggesting only a **minimal rate of change**.
2. **Very Low R^2 Value (0.04):**
   - The **coefficient of determination (R^2 = 0.04)** indicates that **only 4% of the variation** in relative humidity is explained by AQI.
   - This signifies a **very weak linear relationship**, implying that other factors (e.g., weather systems, temperature, wind) have a far greater influence on humidity than air pollution levels.
3. **Data Distribution:**
   - The majority of data points are **clustered horizontally between 70% and 100% humidity**, across a broad range of AQI values (from ~25 to ~175).
   - This concentration suggests that **high humidity is consistently present**, irrespective of AQI.
4. **Color Gradient – Deviation from Fit:**
   - Deviation values are visualized through color, with **most points showing minimal deviation** (dark colors), though some yellow and orange outliers exist.
   - These outliers suggest **occasional mismatches** where observed humidity deviated significantly from the predicted value by the regression line.
5. **1:1 Reference Line (Dashed):**

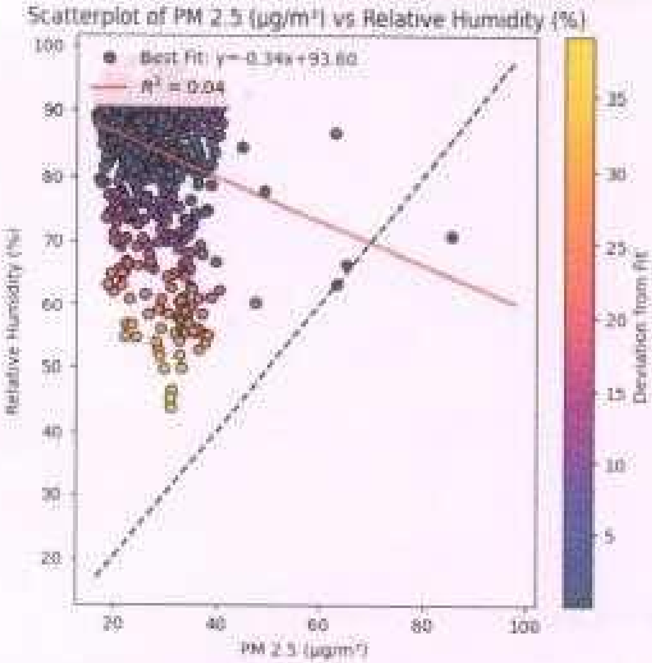


○ The black dashed line serves as a reference but does not reflect any actual relationship here—it's simply a guide.
○ The regression line lies far from it, emphasizing the **non-equivalence** of AQI and humidity values.
Regression equation: Relative Humidity = −0.18 × AQI + 92.57
• **R^2 = 0.04**, indicating **no meaningful linear relationship**.
• Slight **negative trend**, but **humidity remains mostly stable** across AQI levels.
• **Humidity values cluster tightly** around high values (70%–100%).
• Outliers exist, but overall, **AQI and humidity behave independently**.

This scatterplot visualizes the relationship between **PM 2.5 concentration** (fine particulate matter) and **relative humidity (%),** overlaid with a regression line and color-coded deviations to illustrate the fit quality.

1. **Negative Linear Relationship:**
   - **Best-fit regression equation**:
     Relative Humidity=−0.34×PM2.5+93.60\text{Relative Humidity} = -0.34 \times \text{PM2.5} + 93.60Relative Humidity=−0.34×PM2.5+93.60
   - The slope of −0.34 indicates a **negative association**—as PM 2.5 concentration increases, relative humidity tends to **decrease slightly**.
   - However, this trend is **not steep**, meaning changes in humidity are **not very sensitive** to PM 2.5 variation.
2. **Low R^2 Value (0.04):**
   - The coefficient of determination (R^2 = 0.04) implies that **only 4% of the variance** in relative humidity can be explained by PM 2.5 values.
   - This highlights a **very weak correlation**, suggesting that **PM 2.5 levels do not reliably predict humidity**.
3. **Clustered Data Distribution:**
   - Data points are densely concentrated in a narrow PM 2.5 range (roughly between 15 and 45 µg/m³) and high relative humidity values (around 70–95%).
   - This clustering suggests that **high humidity coexists with a wide range of PM 2.5 values**, indicating **no clear dependency** between the variables.
4. **Deviation Color Gradient:**
   - The color bar represents the **degree of deviation from the regression fit**, with warm colors (yellow/orange) showing larger deviations.
   - **Most data points show low to moderate deviation**, but the presence of some outliers indicates variability not captured by a linear model.
5. **1:1 Line (Black Dashed):**

- The diagonal black dashed line represents a 1:1 reference which does **not apply directly** to this comparison, since PM 2.5 and RH are on different physical scales.
- It helps **visually assess divergence**, but has **no direct analytical relevance** here.



- **Equation:** Relative Humidity = −0.34 × PM2.5 + 93.60
- **R^2 = 0.04** → indicates **very weak negative linear correlation**.
- Humidity values mostly lie **above 70%**, regardless of PM 2.5 concentration.
- **High PM 2.5 does not necessarily mean low humidity** and vice versa.
- **Color-coded deviation** helps identify **points poorly explained** by the linear model.

This scatterplot examines the relationship between **PM 10 concentrations** (coarse particulate matter) and **relative humidity (%)**. The linear regression trendline, coefficient of determination (R^2), and color-coded residuals help assess the correlation strength and pattern.

1. **Negative Linear Trend:**
   - **Best-fit line equation:**
     Relative Humidity=−0.19×PM 10+92.99\text{Relative Humidity} = -0.19 \times \text{PM 10} + 92.99Relative Humidity=−0.19×PM 10+92.99
   - The **negative slope** (−0.19) indicates that relative humidity tends to **decrease slightly** as PM 10 levels increase.
   - However, the slope is **gentle**, suggesting that the **relationship is weak** and not dominant.
2. **R^2 Value = 0.04:**
   - An R^2 of **0.04** means that **only 4% of the variability** in relative humidity is explained by PM 10 concentration.
   - This reflects a **very weak linear correlation**, indicating that PM 10 is **not a strong predictor** of humidity.
3. **Data Distribution:**
   - Most data points are **clustered between 30–70 µg/m³** for PM 10 and **70–100%** for relative humidity.
   - Despite some dispersion at higher PM 10 levels, **humidity remains relatively high**, reinforcing the weak correlation.
4. **Deviation and Color Gradient:**

Key Observations and Interpretation

Strong Positive Correlations:

1. **AQI vs PM 2.5 (r = 0.99)**
2. **AQI vs PM 10 (r = 0.99)**
3. **PM 2.5 vs PM 10 (r = 0.99)**
   - These values indicate **extremely strong positive correlations**, suggesting that **AQI is heavily driven by particulate matter**, especially PM 2.5.
   - Since both PM 2.5 and PM 10 contribute to air pollution, this high interdependence is expected.

Weak Negative Correlation:

4. **Relative Humidity vs PM 2.5 (r = –0.20)**
5. **Relative Humidity vs PM 10 (r = –0.20)**
   - These weak negative correlations imply that **higher humidity slightly reduces particulate matter concentration**, possibly due to **moisture helping particles settle or disperse**.
   - However, the correlation is **too weak to be significant on its own**, which matches earlier scatterplot findings.

Moderate Positive Correlation:

6. **Temperature vs PM 2.5 (r = 0.25)**
7. **Temperature vs PM 10 (r = 0.25)**
   - Indicates a **weak-to-moderate positive relationship**: as temperature increases, PM levels tend to **increase slightly**, possibly due to **photochemical reactions** and **stagnant air during warm conditions**.
   - The correlation is weak and may vary with seasonal or local factors.

Strong Negative Correlation:

8. **Relative Humidity vs Temperature (r = –0.84)**
   - A **strong inverse relationship**: as temperature rises, relative humidity tends to fall.
   - This is a **well-known physical relationship**—warm air can hold more moisture, so **relative humidity drops** even if absolute moisture stays constant.
9. **AQI is almost perfectly correlated with both PM 2.5 and PM 10**, highlighting the role of particulate pollution in air quality assessments.
10. **Humidity is weakly inversely correlated with PM**, implying a minor cleansing effect of moisture on air quality.
11. **Temperature has a slight positive relationship with PM and AQI**, possibly influenced by seasonal or urban heating effects.
12. The **strongest non-particulate correlation** is between **Temperature and Relative Humidity** (r = –0.84), confirming well-established atmospheric behavior.

Conclusion

Based on the correlation matrix analysis, it is evident that particulate matter—specifically PM 2.5 and PM 10—plays a dominant role in determining the Air Quality Index (AQI) at the Bhasa Campus. The extremely high positive correlations between AQI and both PM 2.5 and PM 10 (r ≈ 0.99) clearly indicate that variations in AQI are almost entirely governed by changes in particulate concentrations. Relative humidity shows a weak negative correlation with particulate matter, suggesting that increased moisture in the air may have a mild suppressive effect on pollution levels, possibly due to particle settling. Temperature exhibits a slightly positive relationship with particulate levels and AQI, which may reflect the effects of photochemical smog formation and stagnant air during warmer periods. Notably, a strong negative correlation between temperature and relative humidity (r = –0.84) supports established atmospheric dynamics. Overall, while temperature and humidity exert secondary influences, particulate pollution remains the most critical factor affecting air quality in the region.

Name of the members	Signatures
Dr. Debasmrity Mukherjee (Nodal Officer) Dept. of Geography	Debasmrity Mukherjee 22/7/25
Dr. Bidisha Maitra Sen (Dept. of IFF)	
Dr. Sudip Dasgupta (Dept. of Geography)	Sudip Dasgupta 22/7/25
Sri Debabrata Chanda (Dept. of Geography)	Debabrata Chanda 21/7/25
Dr. Shramana Roy Barman (Dept. of Environmental Science)	Shramana Roy Barman
Dr. Sayanti Kar (Dept. of Environmental Science)	Sayanti Kar 22/7/25

Manas Kabi

Dr. Manas Kabi
Principal Asutosh College