Faculty of Engineering, Technology & Built Environment

Department of Electrical and Electronic Engineering

BEE3033 Data Communication & Networks

Assignment (Jan - Apr 2025)

Date :	**23rd February 2025**	
Student Name :	**Pason Cheam Tung Huay**	**Student ID : 1002163305**
Student Name :	**Ahmed Hisham Abdelmoneam**	**Student ID : 1002161742**
Student Name :	**Ng Shang Ping**	**Student ID : 1002267730**

Assessment	**CLO1**	**CLO2**	**CLO3**	**CLO4**	**CLO5**
	PLO1	**PLO2**	**PLO2**	**PLO2**	**PLO4**
Assignment	-	-	-	√	√
Course Learning Outcome	Upon completion of this module, the student will be able to: **CLO4:** Evaluate different methods of data multiplexing, wireless sensor network, and Internet of Things. **CLO5:** Design an application of the Internet of Things.				
Programme Learning Outcome	**PLO2: *PROBLEM ANALYSIS* -** Identify, formulate, conduct research literature and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences (WK1 to WK4);				

	PLO4: ***INVESTIGATION*** – Conduct investigation of complex engineering problems using research-based knowledge (WK8) and research methods including design of experiments, analysis and interpretation of data, and synthesis of information to provide valid conclusions;

Faculty of Engineering, Technology & Built Environment

Department of Electrical and Electronic Engineering

Assignment Marking Rubric

Criteria	Unsatisfactory	Average	Good	Excellent	Score
Introduction & Literature Review [10 Marks]	0 - 2	3 - 5	6 - 8	9 - 10	
	Weak explanation on introduction and literature review.	Appropriately structured introduction and literature review.	Very well explained introduction and moderate enhancement in the literature review.	Excellent explanation on introduction and strong enhancement in the literature review related to the additional components.	
Simulation Work [25 Marks]	0 - 6	7 - 12	13 - 18	19 - 25	
	Irrelevant simulation work.	Only partial simulation work completed.	Simulation work completed, but not methodically.	Simulation work completed methodically, with clear explanation and purposes for every steps.	
Results [25 Marks]	0 - 6	7 - 12	13 - 18	19 - 25	
	No result, or irrelevant results.	Only partial result shown.	Results are presented, but not well organized.	All results are presented in very well-organized manner with clear explanation.	
Discussion [20 Marks]	0 - 5	6 - 10	11 - 15	16 - 20	
	Unable to explain the design results and no analysis.	Able to explain the design results but no analysis.	Able to explain some of the design results with analysis.	Able to explain all the design results and provide analysis with supported results.	
Conclusion [5 Marks]	0	1 - 2	3 - 4	5	
	No conclusion.	Average conclusion of the work done.	Good conclusion of the work done.	Excellent conclusion of the work done.	

Written Skills, Format and References [5 Marks]	0	1 - 2	3 - 4	5	
	Poor in report writing and no references included.	Moderate in report writing, with outdated references. Inconsistent formatting.	Good in report writing and satisfactory referencing has been carried out. Consistent formatting.	Excellent in report writing with correct grammar and provide useful and relevant references. Consistent formatting.	
Presentation [10 Marks]	0 - 2	3 - 5	6 - 8	9 - 10	
	Did not present.	Presentation unconvincing. Able to answer < 50% of questions posed.	Well groomed. Appropriately presented key points. Able to answer > 50% of questions posed.	Well groomed. Convincing in presentation. Appropriately presented key points. Able to answer all questions posed.	
Total Marks [100 marks]					
Final Score (%) = (Total Marks / 100) x 30					

List of Figures

List of Tables

Table of contents

1.0 Introduction

The technology of the Internet of Things (IoT) has significantly transformed how data is collected and used in this modern era. IoT refers to the network of interconnected devices that communicate and share data over the internet, enabling remote monitoring and control of systems in real-time (Kalyesubula, 2018). In order to collect data, the most fundamental requirement are sensors, which are able to detect various aspects in the real world and translate it to signals that can be processed and analyzed. In IoT, data collected from sensors are usually transmitted to the cloud, where it can be organized, stored, and analyzed (Pelaez, 2024). Hence for this assignment, ThingSpeak is used as it provides an open API read and write key for managing sensor data, thus enabling the monitoring of environmental conditions remotely, as well as displaying the collected data through channels in their web application.

Additionally, the ThingSpeak platform is used together with Matlab. The use of sensors with the ThingSpeak platform and processing software like MatLab allows for data analytics which helps to derive insights from the data collected. For this assignment, data are read from an external public channel and written onto our own channel. Matlab Analysis tool is used to perform the read and write action from other public channels to our own channel. Other than that, it is also used for implementing mathematical analysis on the obtained data, performing mathematical processing and computations on the collected dataset to gain in-depth insights into data trends collect from the sensor. Additionally, MatLab Visualization tool is used to present analyzed results in a graphical and informative format. This enables the creation of various types of plots, including **l**ine charts, bar graphs, scatter plots, and heatmaps, which help illustrate sensor data behavior over time.

As the technology of IoT evolves rapidly, its application shows the importance of efficient data collection and real-time analysis for improving critical decision-making moments. IoT represents an important part in data communication and networks, making sure data and information are transmitted properly while maintaining the integrity of the transmitted data through various mediums, architectures, signal modulation and encoding techniques.

Literature Review

Research on the data collection in ThingSpeak are crucial for our assignment, as we will be designing an IoT analytic project with real-time data collection and analysis using ThingSpeak and MATLAB code.

Collecting data are the initial steps towards identifying, evaluating and analysing the cases and problems. There are two approaches towards collecting data. The first approach is the literature study. Literature study is a search and learning stages from various sources such as textbooks and journal related to the research topic (Jamil et al., 2018).

The second approach is the system analysis, this approach aims to analyse the problem obtained from the identification. There are three design stages in this system. The initial stages are the preprocessing phase, it tracks the references related to the research being made and consist of literature studies and system analysis. The design stage and the proposed system are the phase designer and fabrication, this stage contains the design of hardware and the design software. The last phase is the phase testing, this phase will be a process of testing the system created before the implementation (Jamil et al., 2018).

The Internet of Things are technologies connected with various data collecting terminals through internet networks and communication networks. The information of the environment around the object is retrieved in real time, then it converts into a proper data format to be transmitted to the data centre for an intelligent processor to process using cloud computing and various smart computing technologies that can process large set of data (Qiang et al., 2008).

An interface that serves as the data collector that collects data from node devices and also allow for data to be retrieved into the software environment for data analysis are the ThingSpeak platform. ThingSpeak is an opensource application programming interface for storing and retrieving data from the internet or local area network (Saari., 2016).

The stage of system planning can begin from the design of hardware and then follow with the design of software tools. There is a lot of different way to plan the system, but for this research we will be studying in the general system. The system planned are using a sensor to detects any physical or chemical changes in the environment, then convert a certain quantity into analogue units, so that the electronic circuit can read it later. The sensor circuits then transfer to any input / output port for processing. The Arduino will then process the data and then display via personal computer. Lastly, using ThingSpeak to display the data received, this can help the user monitor the various changes received by the sensor circuit (Jamil et al., 2018).

ThingSpeak will be the go-to for public data collection, as it provides real-time data for analyst to perform analysis. And is also a platform that allow researchers to perform analysis that require a lot of data that needs to be keep track of. This study has given us more insight on the data collection process and how ThingSpeak serves as a software bridge for the hardware component such as the sensors and the node devices.

2.0 Simulation Work

In this assignment, we are tasked to design an Internet of Things analytic project that collects and analyse real-time data using public channels from ThingSpeak and MATLAB code like Matlab analytics and visualizations. For our data collection, we decided to collect from 2 public channels on different air pollution levels to interpretate the meaning of the collected data as well as making a comparison of the air quality between two distinct locations. This will allow us to identify different trends and interpretations based on two different set of air pollution data,

Below are the information on the 2 public channels for data collection:

<u>Channel 1</u>
Author: kburghardt
Author Channel ID: 343018

Description:
The channel is called the *"Burghardt's weather station PM10 and PM2.5 sensors with Air Quality Index (AQI)"* by kburghardt. This channel extracts data from a public channel on the amateur air pollution monitoring station in Skierniewice, Poland. Air pollution monitoring channel updates with PM10, PM2.5 and PM1 readings from LookO2, a Plantower PMS5003 based sensor and computed AQI available.

We chose this channel because this channel has been running for 7 years and is still ongoing, while providing detailed information on air pollution.

Link to the kburghardt's ThingSpeak Channel:
https://thingspeak.mathworks.com/channels/343018

<u>Channel 2</u>
Author: angiesoft
Author Channel ID: 820214

Description:
The channel is called *"FeinstaubGreding"* or Fine dust in Greding in English by angiesoft. This channel collects data on fine dust measurement PM10 & PM2.5 with webcam around Greding, Germany. Air pollution monitoring channel also updates with PM10 and PM2.5 readings with computed AQI available as well.

We chose this channel because this channel has been running for a long time and still continuing while providing detailed information on air pollution.

Link to the angiesoft ThingSpeak Channel:
https://thingspeak.mathworks.com/channels/820214

The table below shows the details of our own ThingSpeak channels that extracts data from the public channels mentioned above.

Table 2. 1: Shows the details of our own ThingSpeak channels that extracts data from the public channels mentioned above

Channel	Source channel ID	Our channel ID	Our channel link
1	343018	2851826	https://thingspeak.mathworks.com/channels/2851826
2	820214	2852320	https://thingspeak.mathworks.com/channels/2852320

Data collection

Below are the data collected from both channels.

- **PM10 -** Particulate matter with a diameter of 10 micrometers, including dust, pollen, and mold, which can cause respiratory irritation.
- **PM2.5 -** Finer particulate matter with a diameter of 2.5 micrometers, originating from smoke, vehicle emissions, and industrial processes, capable of penetrating deep into the lungs.

From channel 1 only.

- **Air Quality Index (AQI) -** A standardized measure that assesses air pollution levels based on pollutants like PM10, PM2.5, ozone, and carbon monoxide.

Next, below are the information obtained by performing mathematical analysis with the collected data .

- National Air Quality Index PM10 (NAQI PM10)
- National Air Quality Index PM2.5 (NAQI PM2.5)
- PM 10 Correlation
- PM 2.5 Correlation

Steps for simulation work

To collect and analyze real-time data, we first need to create a ThingSpeak channel. Next, we use the MATLAB Analysis app to process the collected data. Once the analysis is complete, we configure MATLAB Time Control to schedule recurring data collection from the reference channel. Finally, we utilize MATLAB Visualization to generate graphical representations of the analyzed data, providing clear insights.

2.1 MATLAB Analysis App Configuration

1) Read and Write Channel ID

The first step towards data collection is to read the channel ID that we want to collect and then write it to our channel ID, and then we choose the data we need from the referred public channel as shown below in MATLAB code.

```
readChId = 343018;
writeChId = 2851826;  % Replace with your channel number
writeKey = 'P9YEFQ25J9G9Z8GT'; % Replace with your channel write key
[PM10Reading, time] = thingSpeakRead(readChId,'Fields',1,'NumPoints',9);
PM2point5Reading= thingSpeakRead(readChId, 'Fields',3,'NumPoints',9);
```

Figure 2. 1: Shows the read and write channel ID, and read data from the referred public channel

2) Perform Calculation to get the data we need for our channel analysis

After choosing the data needed, calculation is then performed to get the National Air Quality Indices for particle with diameter of 10 micrometres or smaller, PM10 and particle with diameter of 2.5 micrometres or smaller, PM2.5. Calculation is also performed to get the regression equations for PM10 and PM2.5 correlation analysis as shown in MATLAB code below.

```
% Calculation for the  National Air Quality Indices (NAQI)
NAQI_PM10 = (PM10Reading * 100) / 150; % NAQI for PM10
NAQI_PM2_5 = (PM2point5Reading * 100) / 25; % NAQI for PM2.5

% % Calculation for the regresstion equations for PM10 and PM2.5 correlation analysis
PM10_Correlation = 6.0 + 1.1 * PM2point5Reading;
PM2_5_10_Correlation = 6.0 + 0.1 * PM2point5Reading;
```

Figure 1: Shows the calculation for National Air Quality Indices and the regression equation

3) Write all the data into our public channel

After calculation is done, we write all the data we read and calculate back to our public channel ID as shown in the figure below.

```
%Write to our public channel
thingSpeakWrite(writeChId, ...
    [PM10Reading, PM2point5Reading, PM1Reading, airIndexQuality, NAQI_PM10, NAQI_PM2_5, PM10_Correlation, PM2_5_10_Correlation], ...
    'Fields',[1,2,3,4,5,6,7,8],...
'TimeStamps',time,'Writekey',writeKey);
```

Figure 2: Shows the data has been write to our public channel ID

2.2 TimeControl App Configuration

1) Setting up the recurring time controls

For our configuration of the time control app, we set the frequency as recurring as we need to constantly collect the data for minimum of 2 weeks. Then the recurrence was set to every hour, as we will be run the schedule code for every hour to read and write the data into our public channel. The configuration were shown below using ThingSpeak configuration menu.

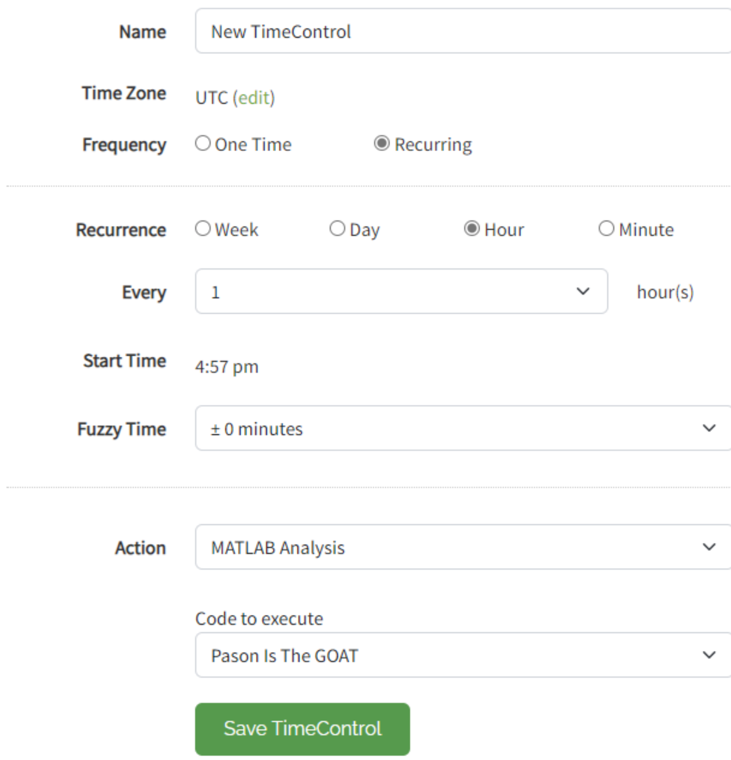

Figure 2. 2: Shows the configuration for the TimeControl App in ThingSpeak

2.3 MATLAB Visualisation App Configuration

1) Read and collect number of points from our public channel

Read our public channel, and then we collect the number of points from our channel to get the correlation of PM10 and PM2.5 as shown in figure below using MATLAB code.

```
readChId = 2851826; %Write your own channel ID
readKey = '7T8YE0V42256KPZX'; %Write your own read key (not write key!)

%collect the number of points from the channel
%[PM10_Correlation, PM2_5, PM2_5_10_Correlation, timeStamps] = thingSpeakRead(readChId,'fields',[6,2,7,1]);
PM10_Correlation = thingSpeakRead(readChId,'fields',7,'NumPoints', 60);
PM2_5 = thingSpeakRead(readChId,'fields',2,'NumPoints', 60);
PM2_5_10_Correlation = thingSpeakRead(readChId,'fields',8,'NumPoints', 60);
```

Figure 2. 3: Shows the MATLAB code of read and collect number of points from our public channel

2) Plot the Scatter Plot

After collect the number of points from our public channel, we then plot the scatter plot of both PM10 and PM2.5 with proper label and legends to show in our public channel page. The MATLAB code to perform this task is shown in figure below.

```
% Scatter Plot
figure;
hold on;
scatter(PM2_5, PM10_Correlation, 'ko', 'filled'); % PM10 (Black circles)
scatter(PM2_5, PM2_5_10_Correlation, 'ko', 'LineWidth', 1.5); % PM2.5-10 (White circles)
plot(PM2_5, PM10_Correlation, 'k--', 'LineWidth', 1.5); % Regression line for PM10
plot(PM2_5, PM2_5_10_Correlation, 'k:', 'LineWidth', 1.5); % Regression line for PM2.5-10

% Labels and Legend
xlabel('PM2.5 concentration (\mug/m^3)'); %Label the X axis
ylabel(sprintf('PM10 or PM2.5-10 \nconcentration (\x03BCg/m^3)')); %Label the Y axis
legend('PM10', 'PM2.5-10', 'PM10 Fit', 'PM2.5-10 Fit', 'Location', 'NorthWest'); %Label the legend
title('PM10, PM2.5, and PM2.5-10 Concentrations'); %Label the title
grid on; %Shows the grid
hold off;
```

Figure 2. 4: Shows the MATLAB code of Scatter plot, Labels and legend

2.4 Website presentation

The extracted data from ThingSpeak are then embedded in a website per the requirement of the assignment. The website is developed manually using HTML CSS and Javascript and hosted using Github.

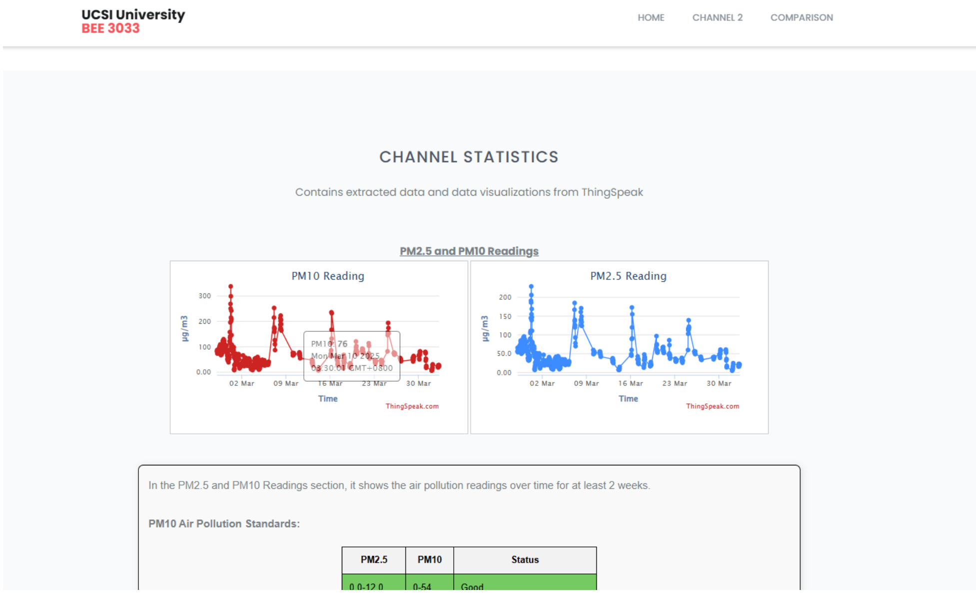

Figure 2. 5: Shows a screenshot of the website to display the ThingSpeak data

Link of the Website:
https://pasoncheam.github.io/Group1DataComAssignment/

2.5 Matlab Codes

Below are the codes for both channels.

Channel 1

Matlab analysis:

```
readChId = 343018;
writeChId = 2851826;  % Replace with your channel number
writeKey = 'P9YEFQ25J9G9Z8GT'; % Replace with your channel write key
[PM10Reading, time] = thingSpeakRead(readChId,'Fields',1,'NumPoints',9);
PM2point5Reading= thingSpeakRead(readChId, 'Fields',3,'NumPoints',9);
PM1Reading = thingSpeakRead(readChId, 'Fields',5,'NumPoints',9);
airIndexQuality = thingSpeakRead(readChId, 'Fields',7,'NumPoints',9);

% Calculation for the  National Air Quality Indices (NAQI)
NAQI_PM10 = (PM10Reading * 100) / 150; % NAQI for PM10
NAQI_PM2_5 = (PM2point5Reading * 100) / 25; % NAQI for PM2.5

% % Calculation for the regresstion equations for PM10 and PM2.5 correlation analysis
PM10_Correlation = 6.0 + 1.1 * PM2point5Reading;
PM2_5_10_Correlation = 6.0 + 0.1 * PM2point5Reading;

%Write to our public channel
thingSpeakWrite(writeChId, ...
    [PM10Reading, PM2point5Reading, PM1Reading, airIndexQuality, NAQI_PM10, NAQI_PM2_5, PM10_Correlation, PM
    'Fields',[1,2,3,4,5,6,7,8],...
'TimeStamps',time,'Writekey',writeKey);
```

Figure 2. 6: Shows the Matlab Analysis for channel 1

Matlab visualizations:

```
readChId = 2851826; %Write your own channel ID
readKey = '7T8YE0V42256KPZX'; %Write your own read key (not write key!)

%collect the number of points from the channel
%[PM10_Correlation, PM2_5, PM2_5_10_Correlation, timeStamps] = thingSpeakRead(readChId,'fields',[6,2,7,1]);
PM10_Correlation = thingSpeakRead(readChId,'fields',7,'NumPoints', 60);
PM2_5 = thingSpeakRead(readChId,'fields',2,'NumPoints', 60);
PM2_5_10_Correlation = thingSpeakRead(readChId,'fields',8,'NumPoints', 60);

% Scatter Plot
figure;
hold on;
scatter(PM2_5, PM10_Correlation, 'ko', 'filled'); % PM10 (Black circles)
scatter(PM2_5, PM2_5_10_Correlation, 'ko', 'LineWidth', 1.5); % PM2.5-10 (White circles)
plot(PM2_5, PM10_Correlation, 'k--', 'LineWidth', 1.5); % Regression line for PM10
plot(PM2_5, PM2_5_10_Correlation, 'k:', 'LineWidth', 1.5); % Regression line for PM2.5-10

% Labels and Legend
xlabel('PM2.5 concentration (\mug/m^3)');
ylabel(sprintf('PM10 or PM2.5-10 \nconcentration (\x03BCg/m^3)'));
legend('PM10', 'PM2.5-10', 'PM10 Fit', 'PM2.5-10 Fit', 'Location', 'NorthWest');
title('PM10, PM2.5, and PM2.5-10 Concentrations');
grid on;
hold off;
```

Figure 2. 7: Shows the Matlab Visualizations for channel 1

Channel 2

Matlab analysis:

```
readChId = 820214;
writeChId = 2852320;  % Replace with your channel number
writeKey = '96SNQG1E866D3EA3'; % Replace with your channel write key
[PM10Reading, time] = thingSpeakRead(readChId,'Fields',1,'NumPoints',9);
PM2point5Reading= thingSpeakRead(readChId, 'Fields',2,'NumPoints',9);

% Calculation for the  National Air Quality Indices (NAQI)
 NAQI_PM10 = (PM10Reading * 100) / 150; % NAQI for PM10
 NAQI_PM2_5 = (PM2point5Reading * 100) / 25; % NAQI for PM2.5

% % Calculation for the regresstion equations for PM10 and PM2.5 correlation analysis
PM10_Correlation = 6.0 + 1.1 * PM2point5Reading;
PM2_5_10_Correlation = 6.0 + 0.1 * PM2point5Reading;

%Write to our public channel
thingSpeakWrite(writeChId, ...
    [PM10Reading, PM2point5Reading, NAQI_PM10, NAQI_PM2_5, PM10_Correlation, PM2_5_10_Correlation], ...
    'Fields',[1,2,3,4,5,6],...
'TimeStamps',time,'Writekey',writeKey);
```

Figure 2. 8:Shows the Matlab Analysis for channel 2

Matlab visualizations:

```
readChId = 2852320; %Write your own channel ID
readKey = '2VH7A6Q0KCK0AXKN'; %Write your own read key (not write key!)

%collect the number of points from the channel
%[PM10_Correlation, PM2_5, PM2_5_10_Correlation, timeStamps] = thingSpeakRead(readChId,'fields',[6,2,7,1]);
PM10_Correlation = thingSpeakRead(readChId,'fields',5,'NumPoints', 60);
PM2_5 = thingSpeakRead(readChId,'fields',2,'NumPoints', 60);
PM2_5_10_Correlation = thingSpeakRead(readChId,'fields',6,'NumPoints', 60);

% Scatter Plot
figure;
hold on;
scatter(PM2_5, PM10_Correlation, 'ko', 'filled'); % PM10 (Black circles)
scatter(PM2_5, PM2_5_10_Correlation, 'ko', 'LineWidth', 1.5); % PM2.5-10 (White circles)
plot(PM2_5, PM10_Correlation, 'k--', 'LineWidth', 1.5); % Regression line for PM10
plot(PM2_5, PM2_5_10_Correlation, 'k:', 'LineWidth', 1.5); % Regression line for PM2.5-10

% Labels and Legend
xlabel('PM2.5 concentration (\mug/m^3)');
ylabel(sprintf('PM10 or PM2.5-10 \nconcentration (\x03BCg/m^3)'));
legend('PM10', 'PM2.5-10', 'PM10 Fit', 'PM2.5-10 Fit', 'Location', 'NorthWest');
title('PM10, PM2.5, and PM2.5-10 Concentrations');
grid on;
hold off;
```

Figure 2. 9: Shows the Matlab Visualizations for channel 2

3.0 Results

3.1 PM10 and PM2.5 Reading

Channel 1

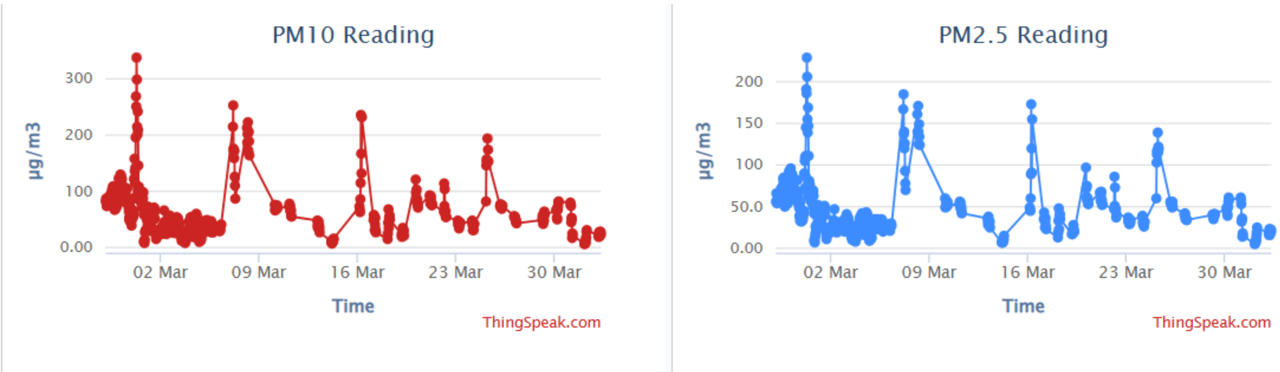

Figure 3. 1: Shows the PM10 and PM2.5 result from channel 1

Based on the graphs, both PM10 and PM2.5 levels show fluctuations over time. The various peaks indicates a significant increase in air pollution on that specific day. The pollution levels decrease after the peak but remains variable. The peaks in the graphs could be due to specific factors like increased vehicle emissions, industrial activity, or environmental factors such as weather conditions or wildfires.

Both pollutants exhibit a similar pattern, with peaks occurring around the same time. This suggests that the sources contributing to these pollutants might be related. The PM10 readings are higher in magnitude compared to PM2.5, which is expected since PM10 includes larger particles.

Channel 2

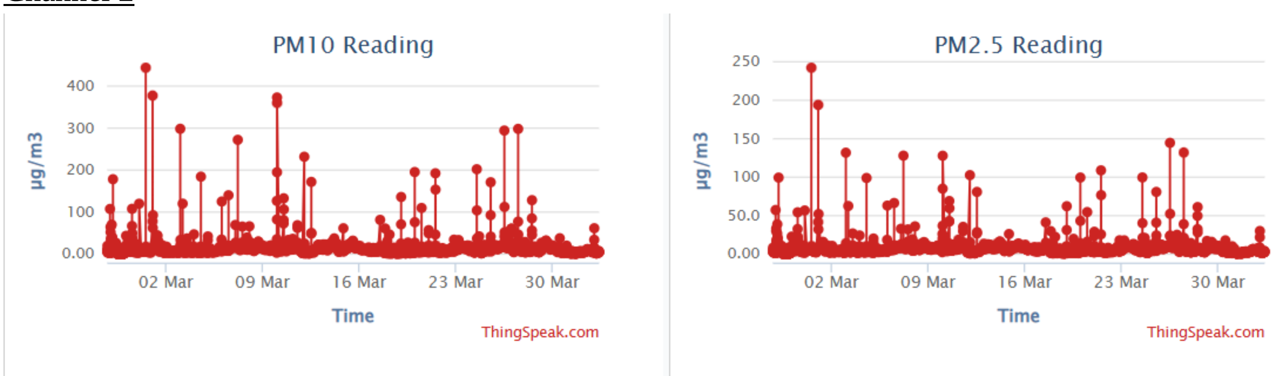

Figure 3. 2: Shows the PM10 and PM2.5 result from channel 2

Based on the graphs, The PM10 readings fluctuate significantly, with occasional spikes exceeding 400 µg/m³. Frequent peaks indicate short-term events of high air pollution. The baseline values seem to stay below 100 µg/m³, but there are multiple sharp increases. Similar to PM10, the PM2.5 values also show fluctuations, though the spikes are lower, reaching around 250 µg/m³ at the highest.

The baseline appears to be lower than that of PM10, staying below 50 μg/m³ most of the time. There are periodic peaks, indicating moments of poor air quality

.

Both PM10 and PM2.5 levels show high variability, with frequent pollution spikes. PM10 generally has higher readings than PM2.5, which is expected as PM10 includes larger particles. The presence of frequent spikes suggests local pollution events, such as construction, vehicular emissions, or environmental factors like haze or dust storms. The general trend suggests that air quality fluctuates but remains poor during certain periods.

Both pollutants exhibit a similar pattern, with peaks occurring around the same time. This suggests that the sources contributing to these pollutants might be related. The PM10 readings are higher in magnitude compared to PM2.5, which is expected since PM10 includes larger particles.

3.2 AQI and NAQI

Mathematical analysis for NAQI

While AQI data is obtained directly from the source channels, the NAQI is calculated by referencing the equation from (Cabello-Torres et al., 2024) shown below.

$$I(PM_{10}) = [PM_{10}] * 100/150$$
$$I(PM_{2.5}) = [PM_{2.5}] * 100/25$$

Formula:

$$I(PM_{10}) = [PM_{10}] \times 100/150$$
$$I(PM_{2.5}) = [PM_{2.5}] \times 100/25$$

Similar to AQI, National Air Quality Index (NAQI) is a measure used to assess and communicate the quality of air based on pollutants, including $PM_{2.5}$ and PM_{10} (Cabello-Torres et al., 2024). However, each NAQI focuses on specific particulate matter.

<u>Channel 1</u>

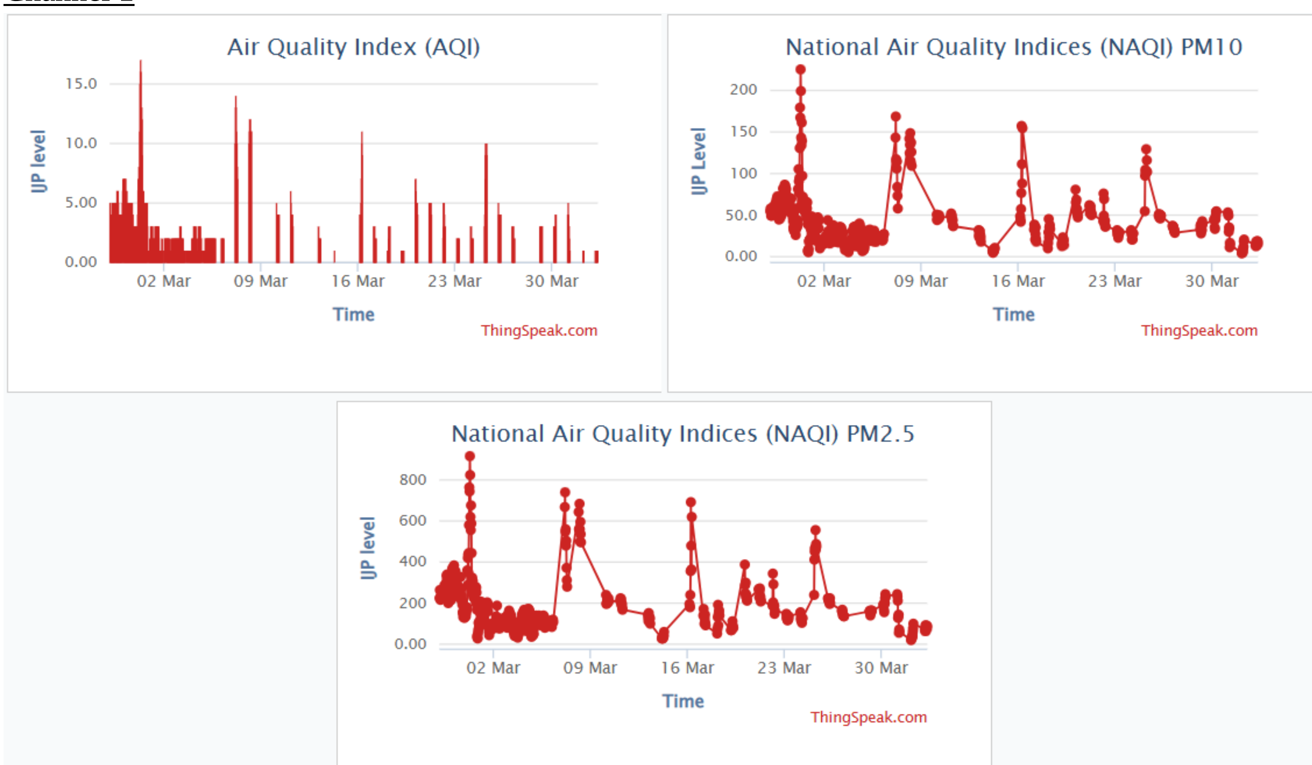

Figure 3. 3: Shows the AQI and NAQI reults from channel 1

The graphs above shows a consistent pattern in air quality trends across the measured metrics. The three graphs also demonstrate a similar peak in pollution levels, indicating 23rd March as a period of heightened air contamination for all pollutants under observation.

Following this peak, a substantial improvement in air quality is also shown, with a clear downward trajectory observed from 23^{rd} March to 30^{th} March. Despite the fluctuations and the mid-month surge in pollution, the overall assessment of air quality during March indicates a generally satisfactory condition within the city. This suggests that while there were specific instances of increased pollution, the prevailing air quality remained within an acceptable range for the majority of the month.

<u>Channel 2</u>

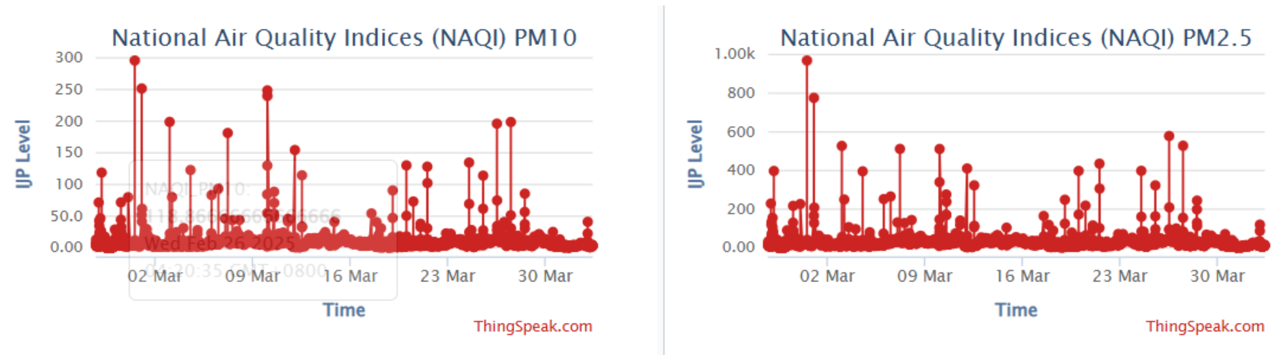

Figure 3. 4: Shows the AQI and NAQI reults from channel 2

Similar to channel one, all three graphs show periodic spikes in pollution levels, especially around the same dates. PM2.5 levels are significantly higher than PM10, highlighting a greater concern for fine particulate pollution. The NAQI values are relatively lower compared to PM indices, but they still exhibit fluctuations. Potential reasons for pollution spikes include weather changes, increased human activities, or industrial emissions.

3.3 PM10 and PM2.5-10 correlation

Matlab analysis for PM10 and Pm2.5 correlation

The correlation for both PM10 and PM2.5 are calculated by referencing the equation from (Janssen et al., 2013).

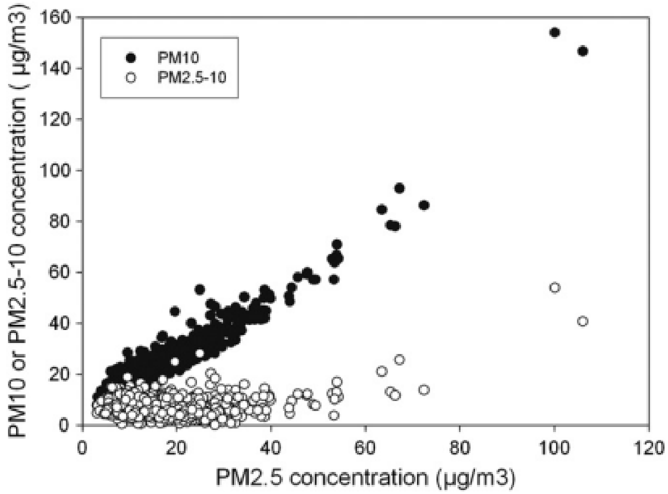

Fig. 3. Relation between PM_{10}, $PM_{2.5}$ and $PM_{2.5-10}$ concentrations. (Regression equations: $PM_{10} = 6.0 + 1.1 \times PM_{2.5}$; $PM_{2.5-10} = 6.0 + 0.1 \times PM_{2.5}$.)

Figure 3. 5: Shows the example graph and equation for the PM correlation from (Janssen et al., 2013)

Formula:

$$PM_{10(correlation)} = 6.0 + 1.1 \times PM_{2.5}$$
$$PM_{2.5\text{-}10(correlation)} = 6.0 + 0.1 \times PM_{2.5}$$

PM correlation refers to the statistical relationship between different types of particulate matter (PM) concentrations measured at different locations or between different PM size fractions (e.g., PM2.5 and PM10). Correlation helps understand how PM levels change together over time and space, which is useful for air pollution monitoring, health studies, and environmental policy-making (Janssen et al., 2013).

Channel 1

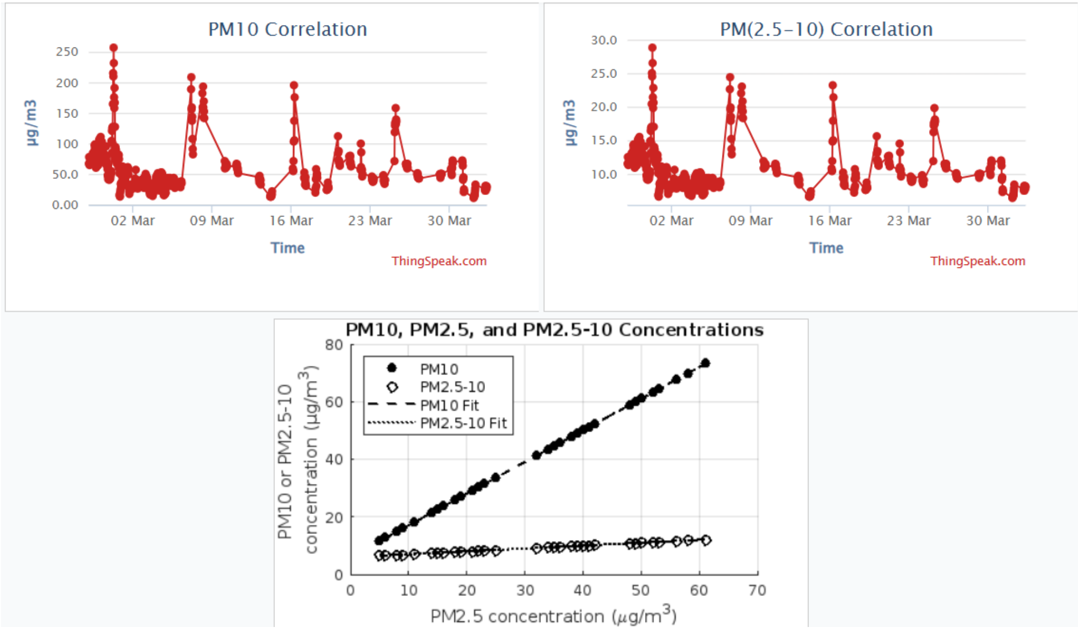

Figure 3. 6Shows the PM correlation result from channel 1

The graphs above indicates variations in PM10 and PM2.5-10 concentrations over March, with noticeable peaks suggesting periodic increases in particulate matter, likely due to environmental factors such as traffic, industrial activity, or weather conditions. The PM10 levels exhibit more significant fluctuations compared to PM2.5-10, which remains relatively lower but follows a similar trend. The correlation plot shows a strong relationship between PM10 and PM2.5, confirming that PM10 largely consists of fine particulate matter, while PM2.5-10 contributes less significantly. These findings suggest that reducing PM2.5 emissions would effectively lower overall PM10 concentrations, highlighting the importance of pollution control measures to improve air quality and public health.

Channel 2

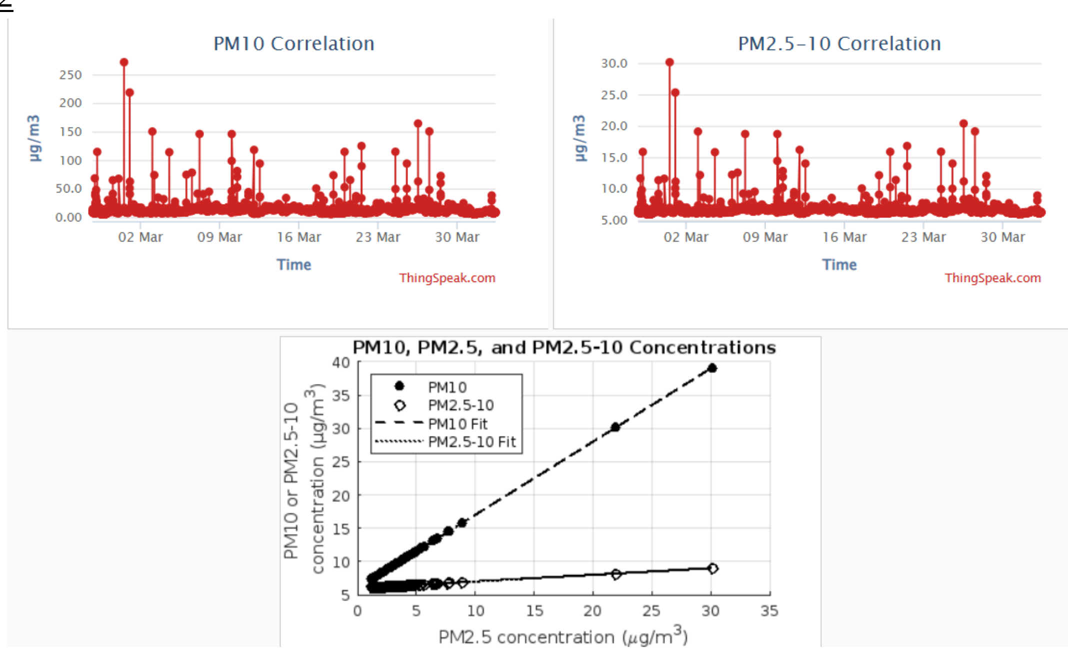

Figure 3. 7: Shows the PM correlation result from channel 2

The graphs above indicates fluctuations in PM10 and PM2.5-10 concentrations over March, with multiple pollution spikes suggesting periodic air quality deterioration. PM10 exhibits larger variations, likely due to external factors like wind and construction, while PM2.5-10 remains more consistent but still shows peaks, posing significant health risks as fine particles can penetrate deep into the lungs. The correlation plot reveals that PM10 is largely composed of PM2.5-10, implying that pollution sources such as vehicle emissions, combustion, and industrial activities play a major role. Given the dominance of fine particulates, controlling PM2.5-10 emissions is crucial for improving air quality and reducing health risks

4.0 Discussion

4.1 Purposes and application of the chosen IoT analytic project.

Purpose

The purpose of this IoT analytic project which extracts data on air pollution on particulate matter, is to collect and analyze real-time air quality data from different geographical locations which are Poland and Germany using public channels on the ThingSpeak platform to monitor air quality trends and generate actionable insights from the collected data. These insights can help develop effective pollution control strategies as well as providing real-life proof for relevant bodies to set proper policies to counter air pollution.

Applications

- **Air pollution monitoring –** By extracting and displaying the data from 2 public channels on ThingSpeak on air pollution, we are able to continuously to monitor the air pollution in both Poland and Germany. This allows us to check pollution levels and fluctuations in air pollution to identity high-risk periods.
- **Obtaining insights into air pollution –** From the extracted data, we can obtain additional insights through mathematical analysis like calculating the NAQI from the data collected to make data-driven decisions. With proper analysis and reasoning, we can provide sufficient information to back up relevant bodies or agencies to develop effective pollution control strategies and policies when the conditions becomes critical.
- **Forecast and predict air pollution –** By analyzing the trend of the data on air pollution, we can anticipate pollution trends and take preventive measures
- **Increase public awareness –** By displaying the extracted data on a website with a user-friendly interface, we can increase the awareness of air pollution of the public by allowing them to view the condition of air pollution over time.
- **Comparison of air pollution conditions –** By extracting data from 2 different countries, we can evaluate the differences in air pollution levels to understand regional variations and contributing factors in air pollution to develop targeted recommendations for improving air quality in the respective countries.

4.2 Actions that can be taken based on your analysis result.

Actions that can be taken:

- **Categorization of air pollution levels**

  From the data collected, it shows massive fluctuations in both channel 1 and 2 with high peaks and low readings. The large range of air pollution readings can be categorized into the following tables.

PM standards:

The table below shows the PM standards on the status of different ranges of PM level. The information fo this table is obtained and verified by (Wambebe & Duan, 2020).

Table 4. 1: Shows the table of the PM standards

PM2.5	PM10	Status
0.0-12.0	0-54	Good
12.1-35.4	55-154	Moderate
35.5-55.4	155-254	Unhealthy for Sensitive Groups
55.5-150.4	255-354	Unhealthy
150.5-250.4	355-424	Very Unhealthy
≥250.5	≥425	Hazardous

AQI standards:

The table below shows the AQI standards on the status of different ranges of AQI level. The information fo this table is obtained and verified by (Karatzas & Kukkonen, 2009).

Table 4. 2: Shows the table of AQI standards

AQI	Status	Description of Air Quality
0-50	Good	Air quality is satisfactory, and air pollution poses little or no risk.
51-100	Moderate	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.
101-150	Unhealthy for Sensitive Groups	Members of sensitive groups may experience health effects. The general public is less likely to be affected.
151-200	Unhealthy	Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects
201-300	Very Unhealthy	Health alert: The risk of health effects is increased for everyone.
≥301	Hazardous	Health warning of emergency conditions: everyone is more likely to be affected.

Correlation table:

The correlation table shows the description of the types of data points and trends, this information is compiled from (Cabello-Torres et al., 2024).

Table 4. 3: Shows the description of the types of data points and trends

Data Points & Trends	Description
Black circles (PM10 data points)	Show a strong linear relationship with PM2.5, indicating a high correlation between the two.
PM10 Fit (Dashed line)	Represents the best-fit line for PM10 data, showing a nearly 1:1 relationship.
White diamonds (PM2.5–10 data points)	These data points are much lower and relatively flat compared to PM10.
PM2.5–10 Fit (Dotted line)	Shows little variation, meaning PM2.5–10 has a weaker correlation with PM2.5.

- **Targeted Pollution Control Measures**

  The identification of pollution peaks and trends in both Channel 1 and Channel 2 indicates the need for targeted pollution control measures. Relevant bodies like the Department of environment (Jabatan Alam Sekitar) can use this data to implement specific interventions, factory and industrial emission controls, or increase public awareness campaigns during periods of high air pollution.

- **Investigation on air pollution sources**

  The correlation analysis between PM10 and PM2.5 in both channels shows that there may be pollution sources as the fluctuations for PM10 and PM2.5 are similar. Hence, further investigation can be taken to find these sources, which may include industrial emissions, vechicle exhaust emmissions or natural events.

- **Comparison between channels**

  While both channels exhibit huge fluctuations in PM10 and PM2.5 levels, the magnitude and frequency of pollution spikes are different from each other. This could be dueto differences in air pollution sources and possible geographical factors.

Below is the table on the comparison between channel 1 and 2.

Table 4. 4: Shows the table on the comparison between channel 1 and 2

Data type	**Differences**	
	Channel 1	**Channel 2**
$PM_{2.5}$	The data shows clusters of high pollution events, where PM2.5 levels increase significantly for extended periods before gradually decreasing. There are noticeable peaks every few days. This suggests that air pollution is influenced by recurring sources or environmental factors such as wind, weather changes, or human activities.	The readings shows frequent short-term spikes but with a more stable baseline pollution level. Channel 2 shows many smaller peaks throughout the time frame, suggesting that air pollution events are more sporadic and temporary.
PM_{10}	Shows high initial pollution levels, exceeding 300 µg/m³ at the start. After the early peak, PM10 levels remain relatively high with periodic surges,	Exhibits frequent sharp spikes, with levels exceeding 400 µg/m³ multiple times. However, baseline pollution is lower, with most data points clustering

	especially around 9th March and 16th March.	below 100 µg/m³. Channel 2 shows short-term pollution events rather than prolonged periods of high PM10 concentration.
NAQI ($PM_{2.5}$)	Large initial spike 2nd March followed by a drop. Several distinct peaks on 9th March, 16th March, and 23rd March, reaching about 600–700 at times. Between peaks, PM2.5 levels remain moderate but persistent, suggesting an ongoing pollution source with periodic intensification.	More frequent but shorter pollution spikes, reaching up to 1,000 in some cases. The baseline remains low between spikes, indicating pollution events rather than continuous pollution. A distinct pattern of sporadic, high-intensity pollution bursts compared to Channel 1.
NAQI (PM_{10})	Initial extreme spike, followed by a significant drop and stabilization at a lower level. Pollution fluctuates, but there are several distinct peaks on 9th March, 16th March, and 23rd March, reaching about 150 at times.	Frequent, sharp spikes reaching 250–300, indicating sudden pollution events. Unlike Channel 1, the baseline PM10 level stays consistently low when no spikes occur. The pattern suggests sporadic high-intensity pollution events rather than sustained pollution.
PM_{10} Correlation	Initial high pollution spike, followed by fluctuating levels. Several peaks on 9th March, 16th March, and 23rd March, reaching about 150–200 µg/m³ at times. Pollution levels fluctuate but remain relatively	More frequent but shorter pollution spikes, occasionally reaching above 200 µg/m³. The baseline remains lower between spikes, suggesting intermittent pollution events rather than sustained pollution. Short, sharp

	high between peaks, indicating a continuous source of pollution.	pollution bursts compared to Channel 1's more sustained pollution levels.
$PM_{2.5\text{-}10}$ Correlation	High initial PM2.5 concentration , followed by fluctuating levels. Several peaks on 9^{th} March, 16^{th} March, and 23^{rd} March, mirroring PM10 trends. Pollution levels remain moderate between peaks, indicating a consistent PM2.5 presence.	Frequent but short pollution spikes, sometimes reaching above 25 µg/m³. Baseline remains lower between peaks, suggesting intermittent pollution bursts. More frequent minor peaks compared to Channel 1, indicating fluctuating pollution events.

4.3 Strengths and shortcomings faced in the project.

Strengths

- The wide variety of air pollution public channels available on ThingSpeak gave a lot of source data choices.
- Strong grasp on web development makes it very easy to create a user-friendly website for data display.
- Having 2 very reliable data extraction source makes it easy to collect data to achieve the project objectives.
- A lot of research on air pollution to refer to in order to interpret the data to provide useful insights.
- Matlab analysis and Matlab Visualization easily allows for the extraction, mathematical analysis and visualization of the data.
- Obtained data source from 2 channel from 2 different countries to make comparative analysis of air quality in two very distinct locations.

Shortcomings

- By using public channels, there is no control are definite knowledge on the sensors used and its accuracy to collect the data. This is overcome by finding channels with more detailed explanation on their channels.
- There is a high chance for the extracted data to be inaccurate due to no control over the physical sensors or sensor malfunctions. This is overcome by finding channels with more detailed explanation on their channels.

- There are a lot of public channels available, but it is difficult to find a reliable one that continues everyday with data that meets the requirements. This is overcome by more searching in ThingSpeak.
- Unfamiliarity with ThingSpeak makes it difficult to get used to extracting data in the beginning of the project. This is overcome by referring to online videos and documentations.
- Unfamiliarity with Matlab Analysis and Visualization takes up time to learn in the beginning of the project. This is overcome through reading the documentations as well as trial and error.
- Difficulty in finding different mathematical equations to perform mathematical analysis duo to the public channels already provide extra details based on the most fundamental PM_{10} and $PM_{2.5}$ readings. This is overcome through more reading of research papers.

4.4 Advantages and Disadvantages of ThinkSpeak in creating an IoT analytic project.

Advantages

- Offers real-time data collection to be kept updated on the latest data trend.
- Utilising IoT in which anyone around the world can access to a public data to see real-time data and perform their own analysis on it.
- Shows the application of data communications and network which demonstrates how IoT devices transmit and exchange data across networks, helping to understand real-world networking concepts.
- Uses cloud-based storage which users can store historical data in the cloud and retrieve it for trend analysis without needing local storage infrastructure.
- Supports a wide range of IoT devices, making it a flexible option for different projects, from environmental monitoring to industrial automation condition checking.
- Has built in software and web application like Matlab Analysis and Matlab Visualization for better mathematical analysis and data visualization.
- Provides a cost-effective platform for IoT projects to store and share real-live data.

Disadvantages

- Utilizing IoT requires dependency in Internet, without stable Internet, data transmission and remote monitoring may be disrupted and unable to view the data in real-live time.
- Since data are extracted from a public channel in ThingSpeak, it is very difficult to determine the accuracy of the sensor used for data collection.
- Utilizing public channels on ThingSpeak also means dependency on the original author for data collection and if the author suddenly stops, there will be no more data for extraction.
- Since data is stored on cloud servers, it may be vulnerable to unauthorized access or breaches.
- Matlab analysis and visualization requires the subscription to Matlab which can be expensive.

- There is no way to inspect the actual place in which the data are collection to verify whether the data collected matches the environmental conditions of the original place from the public channel on ThingSpeak.

5.0 Conclusion

Overall, the objectives of this assignment are achieved successfully, to collect real-time data from ThingSpeak public channel, analyze collected data using proven theoretical mathematic equations, and displaying the analyzed results on a web page. This project utilized IoT to collect and analyze real-time air quality data from 2 public channels on the ThingSpeak platform. The system effectively monitored and compared air pollution levels (PM10 and PM2.5) in two different geographical locations from Poland(Channel 1) and Germany(Channel 2), providing valuable insights into air quality trends and variations in 2 distinct locations.

The use of ThingSpeak and MATLAB facilitated efficient data collection, analysis, and visualization allows for the calculation of NAQI and the identification of correlations between PM10 and PM2.5. These analytical capabilities are crucial for informing public awareness, supporting the development of pollution control strategies, and aiding in decision-making processes by relevant authorities as well as generating crucial insights on the condition of air pollution in those 2 locations.

While the project demonstrated the potential of IoT, data communications and networks for environmental monitoring, it also highlighted limitations, such as the reliance on public data sources and the challenges in ensuring data accuracy and reliability. In the future work, more focus can be put on addressing these shortcomings by incorporating more reliable data sources, enhancing sensor calibration, and expanding the analysis to include a wider range of pollutants.

In a nutshell, this project has significantly enhanced our understanding of IoT-based data extraction, analysis and monitoring. Through this assignment, practical experience in integrating IoT platforms like ThingSpeak with MATLAB are gained, which allows the application of theoretical mathematical equations to real-world air quality data. Additionally, skills in data interpretation, identifying trends, and correlating air pollution parameters across different geographical locations are developed. The project also reinforced knowledge on data communication networks and cloud-based data

6.0 References

1. Kalyesubula, N. (2018). *Challenges in control and monitoring of home-based electronic devices: A case for IoT-based automation system*
2. https://www.academia.edu/75726691/Challenges_In_Control_And_Monitoring_Of_Home_Based_Electronic_Devices_A_Case_For_Iot_Based_Automation_System

3. Pelaez, A. (2024). 10+ Examples of IoT Sensors & How Data Is Used in 2024. https://ubidots.com/blog/iot-sensor-data/

4. Jamil, M., & Said, M. (2018). The utilization of internet of things (IoT) for multi sensor data acquisition using thingspeak. VOLT: *Jurnal Ilmiah Pendidikan Teknik Elektro*, 3(1), 13-18.

5. Qiang, C., Quan, G. R., Yu, B., & Yang, L. (2013). Research on security issues of the internet of things. *International Journal of Future Generation Communication and Networking*, 6(6), 1-10.

6. Saari, M., Baharudin, A. M., Sillberg, P., Rantanen, P., & Soini, J. (2016, May). Embedded Linux controlled sensor network. In 2016 *39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1185-1189). IEEE.

7. Cabello-Torres, R. J., Carbo-Bustinza, N., Romero-Cabello, E. A., Ureta Tolentino, J. M., Torres Armas, E. A., Turpo-Chaparros, J. E., ... & López-Gonzales, J. L. (2024). An exploratory analysis of PM2.5/PM10 ratio during spring 2016-2018 in Metropolitan Lima. *Scientific Reports*, *14*(1), 9285.
8. Janssen, N. A. H., Fischer, P., Marra, M., Ameling, C., & Cassee, F. R. (2013). Short-term effects of PM2.5, PM10 and PM2.5−10 on daily mortality in the Netherlands. *Science of the Total Environment*, *463-464*, 20-26.
9. Wambebe, N. M., & Duan, X. (2020). Air Quality Levels and Health Risk Assessment of Particulate Matters in Abuja Municipal Area, Nigeria. *Atmosphere*, *11*(8), 817.
10. Karatzas, K., & Kukkonen, J. (Eds.). (2009). *Quality of life information services towards a sustainable society for the atmospheric environment*. COST Action ES0602 Workshop Proceedings.