Examining the Air Quality Index and its Adverse Effects: An Analytical Study for smart Cities Development

Dr. Rohit Rastogi*

(Associate. Professor, Dept. of CSE, ABES Engineering College Ghaziabad, U.P., India) rohitrastogi.shantikunj@gmail.com, 8076772048; 9818992772

Priyanshi Garg

(Student, B.Tech. Dept. of CSE, ABES Engineering College, Ghaziabad, U.P., India)

priyanshi.18bcs1068@abes.ac.in, 9911275441

Mukund Rastogi

(Student, B.Tech. Dept. of CSE, ABES Engineering College, Ghaziabad, U.P., India)

mukund.18bcs1021@abes.ac.in, 9628148953

Prajwal Srivastava

(Student, B.Tech. Dept. of CSE, ABES Engineering College, Ghaziabad, U.P., India)

prajwal.18bcs1005@abes.ac.in, 7233039989

Rishabh Jain

(Student, B.Tech. Dept. of ECE, ABES Engineering College, Ghaziabad, U.P., India)

rishabh.17bec1024@abes.ac.in, 9354571216

Puru Jain

(Student, B.Tech. Dept. of ECE, ABES Engineering College, Ghaziabad, U.P., India)

puru.17bec1004@abes.ac.in, 9910625014

Neha Gupta

(Student, B.Tech. Dept. of CSE, ABES Engineering College, Ghaziabad, U.P., India)

mukund.18bcs1021@abes.ac.in, 8368860404

Author's profile

Dr. Rohit Rastogi received his B.E. C. S. S. Univ. Meerut, 2003. Master's degree in CS of NITTTR-Chandigarh from Punjab University. He obtained his doctoral degree from the Dayalbagh Educational Institute in Agra, India. He is an associate professor in the CSE department of ABES Engineering College, Ghaziabad, India. He has won awards in a several of areas, including improved education, significant contributions, human value promotion, and long-term service. He keeps himself engaged in various competition events, activities, webinars, seminars, workshops, projects and various other educational learning forums. He strongly believes that Transformation starts within self.

Ms. Priyanshi Garg is a student of B. Tech (CSE) in ABESEC which is affiliated to AKTU. She is currently working on Yagyopathy where she is analyzing the data and translate them. She has a keen interest in coding and cyber security. Her hobbies are to watch movies. She wishes to do something for her society in coming future with her all resources. She has many research papers published in her credit in international publisher while studying in graduation. She wish to be a successful IT engineer and develop some technical solutions for rural areas. Priyanshi is ambitious and hardworking and she is very sincere and always before time to submit any task assigned. She has team spirit and polite gesture.

Mukund Rastogi is engineering student in AKTU Univ. Presently he is B. Tech. Final Year student of CSE in ABESEC, Ghaziabad. India. He is working presently on Yagya and Mantra therapy and its analysis by Machine Learning. He has keen interest in Google surfing. His hobbies is playing badminton and reading books. He is young, talented and dynamic. He is placed in a good IT company and strong interest in Data sciences. He is versatile and smart personality and wish to serve country through IT sector. He has developed some good analysis for different data science projects.

Prajwal Srivastava is engineering student in AKTU Univ. Presently he is B. Tech. Final Year student of CSE in ABESEC, Ghaziabad. India. He is working presently on Yagya and Mantra therapy and its analysis by Machine Learning. He has done projects on IOT, Nodejs, Reactjs, open CV. He has keen interest in cyber security. His hobbies is to listen music and reading books. He is young, talented and dynamic.

Rishabh Jain is engineering student in AKTU Univ. Presently he is B. Tech. Final Year student of ECE in ABESEC, Ghaziabad. India. He is currently working on Yagya and Mantra therapy and its analysis by Machine Learning. He is interested in field of cyber security. His hobbies is listening to music and reading books. He is young, talented and dynamic.

Puru Jain is engineering student in AKTU Univ. Presently he is B. Tech. Final Year student of ECE in ABESEC, Ghaziabad. India. He is working presently on Yagya and Mantra therapy and its analysis by Machine Learning. His hobbies is playing chess, reading books and listening music. He is young, talented and dynamic.

Neha Gupta is a student of B. Tech (CSE) in ABESEC which is affiliated to AKTU. She is currently working on Yagyopathy where she is analyzing the data and translate them. She has a keen interest in coding and cyber security. Her hobbies are to watch movies. She wishes to do something for her society in coming future with her all resources.

Abstract. Because of quick industrialization and urbanization, Delhi, NCR in India is seeing uncommon increment in air contamination. The arrival of unsafe gases, for example, carbon monoxide and sulfur dioxide aren't just hurtful to the soundness of the populace, but on the other hand is making irreversible impacts nature. The motivation behind this undertaking is to depict the execution of ease and vitality proficient air quality checking frameworks. The proposed air quality observing framework utilizes an air quality record that can be effectively translated. To show the truth of the gases being discharged around parts of Delhi.

The Air Quality Index has been created by the Environmental Protection Agency (EPA), to give every day levels of air contamination and a lot increasingly precise, effectively reasonable data. The Air Quality Index (AQI) is a table created by this Environmental Protection Agency (EPA) to give every day and precise data about air contamination to normal residents. AQI considers four significant poisons: sulfur dioxide, carbon monoxide,

nitrogen dioxide and ozone and benzene, SPM. By recalling information on day by day contamination levels, the EPA can indicate a number about its AQI run. The assessment results affirm the ability of the proposed framework for constant and spatial observing of air quality. Also, it is workable for the overall population to arrive at the aftereffects of air quality observing continuously.

Keywords/ Search items. Environmental Protection Agency, Air Quality Index, PM level, Tropical Meteorology, Expert system, Machine learning, deep learning, Drugs discovery, Big Data Analytic, Healthcare technology, Explainable Artificial Intelligence

1 Introduction

With standing zone and no water, air is the significant resource for living. With inventive progress, an immense proportion of data has been made fusing air quality and is utilized to build up the idea of the air in different regions. The best outcome is to give AQI the alternative to do it proficiently. Delineations of such air quality are ordinarily confused with strategic and particular systems, in any occasion, as the measure of increment in model stations and pollution parameters. Along these lines we need to keep the AQI site all together.

Concerning the general populace, they would not be content with the crude data, time-request plots, quantifiable checks and other complex disclosures identified with air quality and atmosphere. For that and virtualizes AQI. The outcome is that people will regularly lose interest and can't invite air quality conditions nor sullying anticipation endeavors by authoritative workplaces. From this time forward we will perceive and do likewise thusly (Azam, A. G. et al., 2016).

To see the above factor and issues, the possibility of air quality index has been detailed and utilized sufficiently in many fabricated nations in the ongoing three decades. An index for quality of air is really known as a routine schedule that measures and indicates weighted evaluations of individual and air defilement characteristics and indicators (SO2, NO2, SPM, etc.) Into singular numbers or sets of numbers. Outstanding endeavors have been made to make and utilize AQI in India, initially in view of a subtle air quality testing and program explicitly propelled in 1984 and with a receptive outlook about air defilement Practicing was basically nonexistent. Thusly we can do it suitably.

1.1 Health Care

Endless the AI (ML) industry's most sultry youthful new associations are delving in noteworthy bits of their endeavors to social security, including Nervanasys (beginning late got by Intel), Ayasdi ($94MM), Sentient.ai ($144MM), Digital Reasoning Systems ($36MM) (Raj., P. et al., 2016).

With all the eagerness in the scholar and research frameworks, we at Emerj have discovered most AI specialists increase some hard encounters putting a finger on where AI is vehemently shaping restorative organizations today. We've made this article, not to be a finished supply of potential applications, at any rate to remember distinctive present and future employments of AI for the helpful field, with vital interfaces with outside sources and related Emerj interviews (Kasthuri, A., 2018).

1.2 Machine Vision

In the wide extent of AI's present typical desires, AI helpful organizations applications appears top synopsis financing over most recent years.

In mid-2013, IBM utilized in medical and in the wake of dominating a dumbfounding match planning game in against world's best living Go player, chose heave restorative chances advancing also (Chauhan, S. et al., 2017).

1.3 Medical Imaging and Artificial Intelligence

Starting late, man-made cognizance (AI) got area into standard every day presence in various habits, from language affirmation mechanical assemblies on mobile phones to the assessment of cash related trades, to counts for self-driving cars, or for playing the key tabletop game Go. Restorative imaging, also, is likely going to encounter a focal change as soon as possible. "It is definitely not hard to envision that AI will be continuously realized in restorative imaging structures," Italian expert Francesco Sardanelli commented in an article feature on winning examples in radiology. In like manner, as demonstrated by a progressing review, over portion of worldwide therapeutic administrations pioneers foresee the activity of AI in watching and investigation to expand (Brondino, N. et al., 2013).

Despite the way that the use of AI is presently typical practice in specific pieces of the field of imaging, publicize examination predicts a further impact all through the accompanying five to 10 years. Fresher AI methodologies, for instance, "significant learning," could prepare for quantitative, regulated, yet in like manner tweaked imaging, while simultaneously preventing expressive errors and, at the same time, enabling bolstered gainfulness increases. Radiologist Keith Dreyer of Harvard Medical School worried at an expert gathering in the US that "Significant AI will improve quality, viability, and results (Brondino, N. et al., 2013).

1.4 Correlation of Air Pollution and Human Health

Our physical and mental success is affected differently by the kind of air defilement we are exhibited to. There are various organs and genuine limits that can be harmed, the outcomes including:

- Respiratory disorders
- Cardiovascular mischief
- Exhaustion, cerebral torments and uneasiness
- Irritating of the neck, sell and view organs
- Mischief to regenerative organs
- Harm to the digestion system, respiratory and blood system.
- Tangible framework hurt (dash, S. et al., 2019); (Rastogi, R. et al., 2022).

1.5 The Effect of Polluted air on life of Human Species

The brief effects of air sullying are hard to disregard. Watery eyes, inconvenience breathing are extreme and standard reactions.

A normal 92 percent of the absolute people live in locales with dangerous degrees of air pollution and, even at clearly elusive levels, air defilement can extend one's risk of cardiovascular and unexpected passing (Sharma, A. et al., 2019).

1.6 Annals/ Details of Air Pollution in Global Scenario

Air pollution causes 7 million surprising misfortunes consistently, and passing on account of incorporating air tainting are on track to twofold by 2050. Notwithstanding, air defilement is recognized by various assemblies as an unavoidable consequence of financial improvement. Neutralization of air defilement is underfunded in real improvement help and no noteworthy foundation has made air sullying its need (Veni, K. et al., 2014).

Five national establishments of science and medication have now said something in regards to the dangers of air defilement for human prosperity. The science establishments of Brazil, Germany, South Africa, and the USA and the US National Academy of Medicine have assessed the confirmation accomplice particulate air sullying to coronary ailment, stroke, wearisome obstructive lung infection, lung threat, less

than ideal birth, diabetes, and psychological well-being. The Academies consider this to be as unequivocal. They along these lines call for sincere movement to control air sullying and propose apportionment of another Global Compact on Air Pollution and Health (Rastogi, R. et al., 2021).

1.7 Application of DL and AI in Air Pollution and Healthcare

Watching air quality is an indispensable activity in various current and urban zones of our planet. Air Quality Operational Centers (AQOCs) are set up hence, in locales with certifiable air tainting issues. The AQOCs are operational units, at risk for regulating checking frameworks, setting up the accumulated information, over the long haul giving on-line evaluation of air tainting and its present minute and long stretch improvement. Amazing volumes of data are accumulated after some time in these core interests. One of the prime stresses of analysts in AQOCs is the usage of acceptable showing gadgets that award explanation and endorsement of the assembled data. Man-made intelligence techniques have all the earmarks of being sensible in this one of a kind condition (WHO, 2010).

1.8 Big Data Analytics in Healthcare

Huge data in social protection and medication insinuates these diverse tremendous and complex data, which they are difficult to explore and manage with standard programming or gear. Huge data assessment covers coordination of heterogeneous data, data quality control, examination, illustrating, clarification and endorsement. Usage of enormous data assessment gives broad data finding from the open tremendous proportion of data (WHO, 2010).

Particularly, immense data assessment in medicine and social protection enables examination of the colossal datasets from an enormous number of patients, recognizing gatherings and association between's datasets, similarly as making perceptive models using data mining techniques (WHO, 2010).

2 Motivation, Objective and Scope of the Research Work

2.1 Motivation

a- Resource Allocation: To help the chairmen in apportioning the assets and deciding the needs. The ordinary exchanging must be done so as to know the asset. Empower assessment of exchange offs associated with the other air contamination control techniques.

b- Location Based Dashboard: It is used to measure the harmful air quality at different places and its consequences in different angles at developed and developing areas. What's more, along these lines we can rank the areas. Along these lines, bringing up those territories and frequencies of potential risks.

c- Trend Analysis: To reflect the different variations in the standard of air index this has been happening over a specified time. So as for to add the trend and thus simplify it. This enables forecasting of air quality and plans pollution control measures.

d- Public Information: To educate general society about the ecological conditions. It's important for the peoples who have suffered from harmful effects caused by the air contamination. The individuals should think about the data identified with contamination. In this way it so empowers them to adjust their day by day work and imparts at the moment when they are aware regarding higher contamination levels.

e- Scientific Research: It is used as a technique to reduce this data and converts it in human friendly and easy to use structure and shape by preprocessing the data that delivers more understanding and clarity to the researchers while directing the examination over some natural wonders. The science, for example, to be fighting and ready to invalidate it. Such instruments become increasingly helpful when utilized in along these lines the combination with different sources, for example, we realize nearby outflow studies.

2.2 Research Work Objective

The project aims to achieve the following:
Inform all over to the public about the air quality by summation parameter which is easily understandable.

Informing citizens about the overall and every concerned effects of the parameters in healthcare 4.0 of the poor quality of air and climate related concerns.

Developed urban areas are marked points and towns for beautifully given smartly priority on the responses provided to take care the quality of the air.

2.3 Scope of the Project

- Review of the all the available AQIs and also including international practices and exercises.

- Suggest health impact has to be limited above the pre-decided finalized limits for the 8 factors for which the quality protocols of air in small duration are the recommended. The overall the air quality index and overall presentations.

- Suggest qualitative description of and material analysis of the air quality and related same health impressions for all the many other numerical annals related to quality of air.
- The measurement of quality of air with the available data and so overall from a less but already marked developed and developing areas of residential for citizens.

- To improve the IoT and online based and easily manually usable structure for handling quality of air for the citizens applying on present and previous parameter values related to air quality. The web-based project will be promising.

3 Related Previous Work

In the past barely any decades, AQI has been set up on a most fundamental sub-list approach utilizing five parameters for example suspended particulate issue, SO2 CO, PM-ten and NO2. As they say, we need to get from the past works. In any case, the concluded AQI was constantly coordinated by the sub-summary of SPM because of the nonattendance of information transparency for different damages present in the earth. The IITs have been dealing with this issue reliably. Indian Institute of Tropical Meteorology, Pune has been developed an AQI, which gives a report to PM-ten, PM-two pint five, O3, NO2, and CO, and has applied to driving forward air quality watching systems (A. Masih, 2019).

The IITM-AQI depicts air quality in the focal points of entirely heartbreaking, remarkably poor and poor (annoying for sensitive get-togethers), moderate and mind blowing. So in a relentlessly probabilistic way, we can way to deal with oversee have better results (Suliankatchi, R., et al., 2013).

The reconsidered National Ambient Air Quality Standards are provoked for the twelve parameters – PM10, PM2.5, NO2, SO2, Pb, Ni, As, and Benzene. Doubtlessly, in a way we need to add on to the past works. Despite the manner in which that AQI is ordinarily settled on criteria toxic substances and gases, another way to deal with oversee AQI which considers whatever number contaminations from the review of told hurts as could reasonably be typical is beguiling. Regardless, the picking of such parameters on an essential level relies on the AQI objective, information accessibility, averaging periods and the checking rehash, and estimation strategies. While PM10, PM2.5, NO2, and Pb have 24-hourly furthermore the yearly common principles, Ni, SPM, As, and benzene have as of late yearly measures and CO and O3 increase some little encounters of period models. PM10, PM2.5, SO2, NO2, CO, and O3 are settled on the tireless explanation at the many air quality stations (counting NH3 at explicit stations), and NH3, at whatever point checked, use when in doubt manual frameworks. In setting on the above foundation, Central Pollution Control Board (CPCB) has been started and this task on National Air Quality Index to sustain air quality data about the scattering structure for more noteworthy open consideration as needs be

as their assistance declaring continuously conditions the board (Rastogi, R. et al., 2023a).

4 Software Requirement Specification, Product Documents

4.1 Product Perspective

AQI can be extended by growing air surges, for instance, during obstructed hour traffic or when there is a dismissed quickly spreading fire, or by the debilitating of air poisons and gases. The general direct of the thing will depend upon it. Stale air, much of the time realized by an anticyclone and temperature separate or diminished breeze speed, grants air pollution to remain in a constrained region, causing high groupings of poisons, invention reactions between air contaminants and dim conditions is.

The standards and benchmark for parameters for quality of air gives numbers stretching out from one to ten and onwards to exhibit the level of prosperity danger associated with quality of air in nearby region. To a great extent this ought to be conceivable and improved. From time to time, when the proportion of air sullying is uncommonly high, the number can be at any rate 10. AQHI gives a local total of the most outrageous gauge of air quality for a predominant close by air quality present a motivating force similarly with respect to now, today and thusly tomorrow, and gives related prosperity counsel.

In busy time gridlock related sources, vehicles drive attempts to construct numbers and lessen mileage outpourings. Rather than standard reasoning, by far most of the air tainting in Delhi isn't a direct result of vehicular traffic. Progressing assessments show that road dust (half) and industry (23%) - vehicles are simply 7%. Among the cutting edge suppliers, power plants inside Delhi city limits were the rule blameworthy gatherings. As showed by a draft report by IIT Kanpur, road buildup and helper particles from various sources are the fundamental supporters of defilement in the national capital. While sullying is about the equal in both summer and winter, sources said that vehicles are the huge supporters of defilement during winter.

4.2 Interfaces of the Hardware and Software Designed

Hardware Interfaces

PROCESSOR	:	Intel dual Core RAM	:	4.0GB
HARD DISK	:	80GB		

Software Interfaces

OPERATING SYSTEM	:	Windows 7/XP/8/10		
FRONTEND	:	Html, CSS, Javascript SERVERSIDESCRIPT	:	Php
LANGUAGE	:	R		
FRAMEWORK	:	RSTUDIO		

4.3 Site Adaptation Requirements

The AQI framework ought to have been electronic AQI scattering which ought to be made arrangements for online check and show of the nation over AQI. The general working of the site will be portrayed here. The site ought to be rendered a catalyst, basic and a simple looking reaction to the AQI question. Different highlights of this site ought to be related with the general proclaiming of toxic substances and gases committed for the record. Toxins of the beating benchmarks and success impacts the inhabitants and condition.

In spite of the site, we can also render menu-based AQI by methods for looking through states and the urban districts from the dropdown menu. The dependable highlights will be related with the site. As a second piece of the handiness, the site can in like way be rendered for menu-based AQI demand by

methods for looking through bits of Delhi.

4.4 Product Functions

The usefulness will show crafted by AQI and the client can utilize the site to control as indicated by his necessities and experience the information to break down the meteorological forecasts for the future viewpoint. This item will give the future climate projections as per the past information. The last item will be so easy to understand and furthermore effectively hand capable for the simple utilization of the individuals. This item will give better alternatives and completely utilitarian condition.

4.5 User Characteristics

To interpret air quality through Air Quality Index (AQI) and understand health impacts of air pollutant concentration levels monitored. To estimate benefit and cost of air pollution abatement in Delhi using

a) Household health production function

b) Demand function

4.6 Assumptions and Dependencies

According to the AQI system, air quality is associated with the mortality and the respiratory and morbidity rate of the city of Delhi. The dependencies are taken from the previous data. There must be regular collection of the record of density of particles responsible for polluting the air which were interpreted into the AQI values for many various air quality check and control centers for the period of 2011-2015.

1. Air quality monitoring stations were compared on the bases of on yearly percentage trend in each of the health categories and results in assumptions.
2. Higher AQI value denotes poor air quality and majorly the high ratio of the crowd gathering is supposed to get the exposure and skillfully growing highly and severe and ill impressions on the health (Refer Fig. 1, 2 and 3 for Use case, class and architectural diagrams of the product proposed).

4.7 Use case

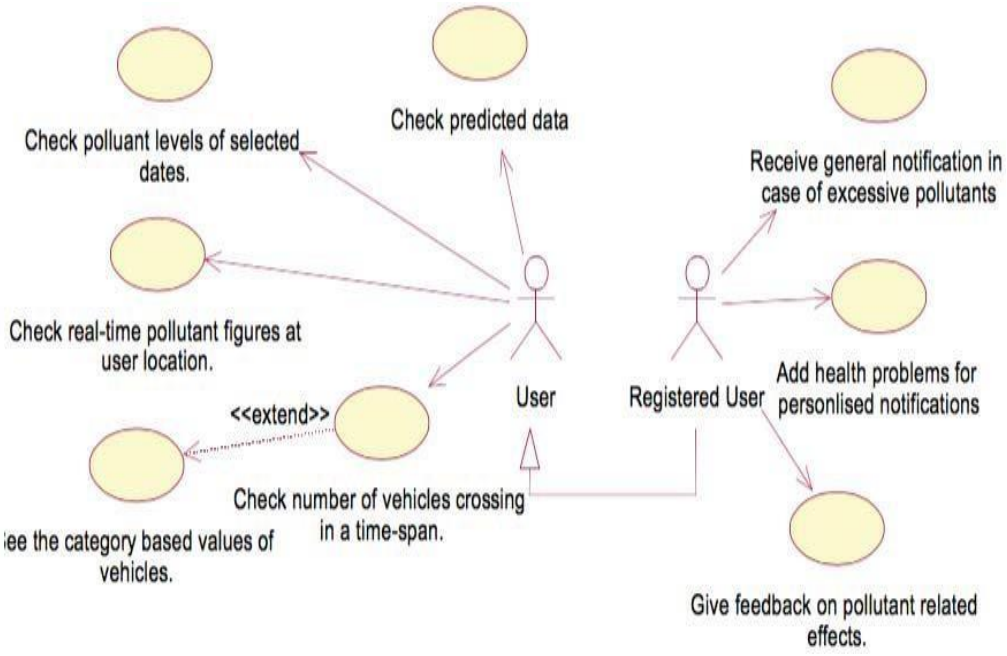

Figure 1. *Use Case for the Web Interface*

5 System Design

5.1 Architecture Diagram

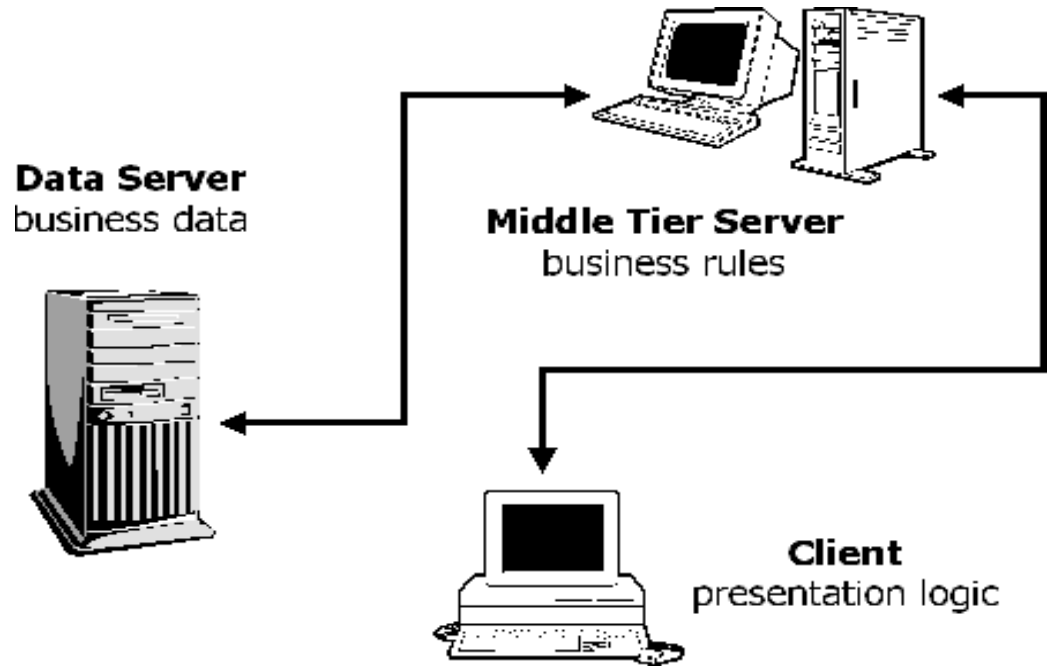

Figure 2. *Tier Architecture Diagram*

5.2 Class Diagram

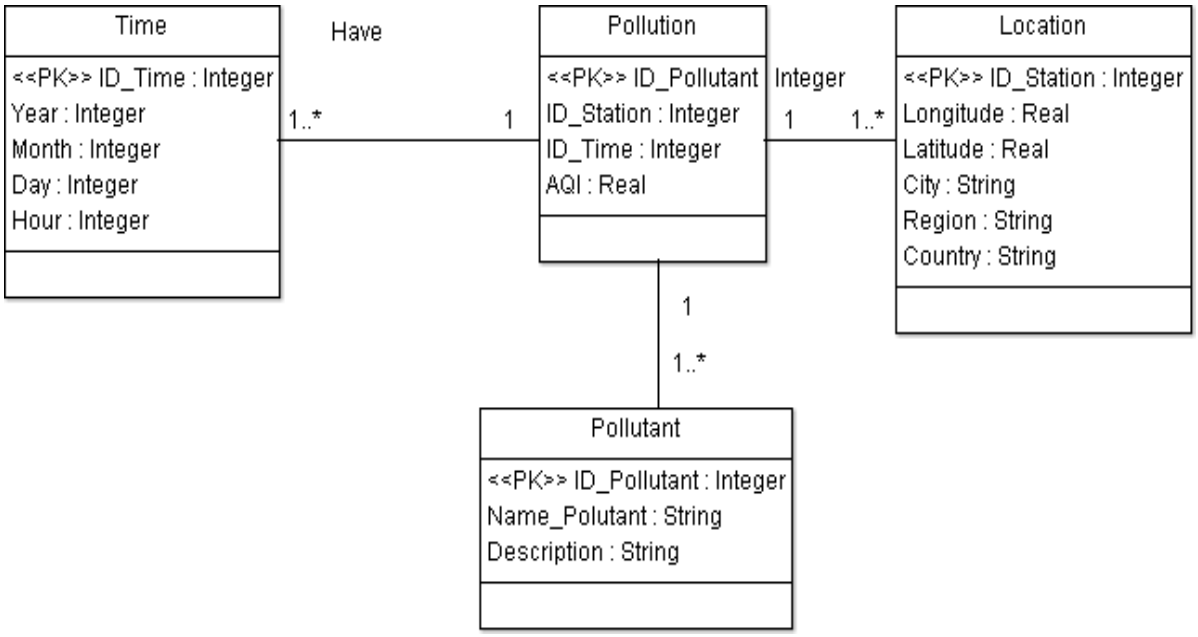

Figure 3. *Class Diagram of the Product Designed*

6 Implementation and Results

6.1 Assumptions and Dependencies

AQI for some particular day and at some perfect territory can be truly dictated by using the MS Excel wherein a straightforward evaluation of AQI has been made. The general endeavors doubts and conditions are trailed by this structure. The customer needs to enter in any occasion three estimations of any pollution center in the blue cells and the sub-records are then decided and examined in this way indicating the remainder of AQI close by the concealing importance the AQI characterization. The prosperity impacts are further identifying with the AQI class are bare essential at the base of the sheet.

By and large AQI is resolved just if data are available for least three poisons out of which one should basically be among those three. The accuracy must be resolved and subsequently do moreover. Else, data is seen as insufficient for figuring AQI. Also, in like manner at any rate in any occasion 16 hours data is seen as significant for figuring sub-record (Ref. Fig. 4, 5 and 6 for snapshot of Analysis of yearly gas emission).

6.2 Implementation Details and Snapshots of Interfaces

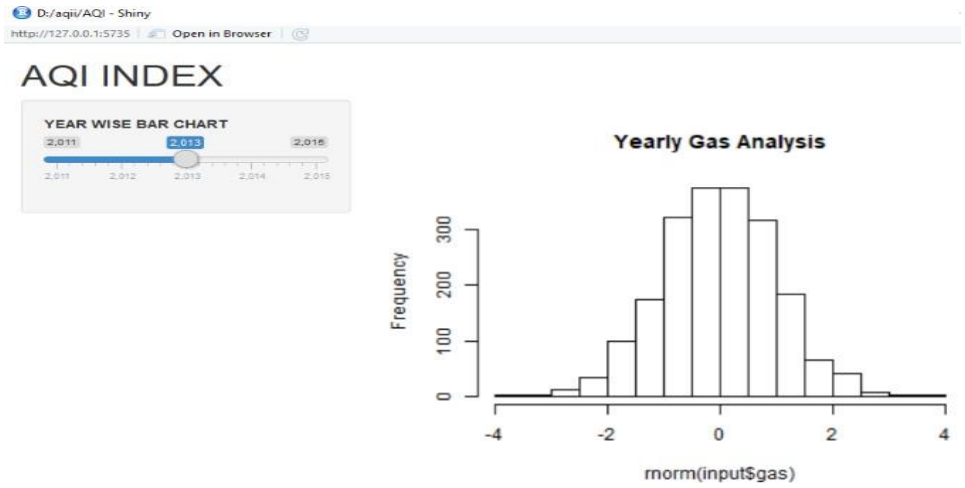

Figure 4. *AQI Index and yearly Gas Analysis (Snap Shot-1)*

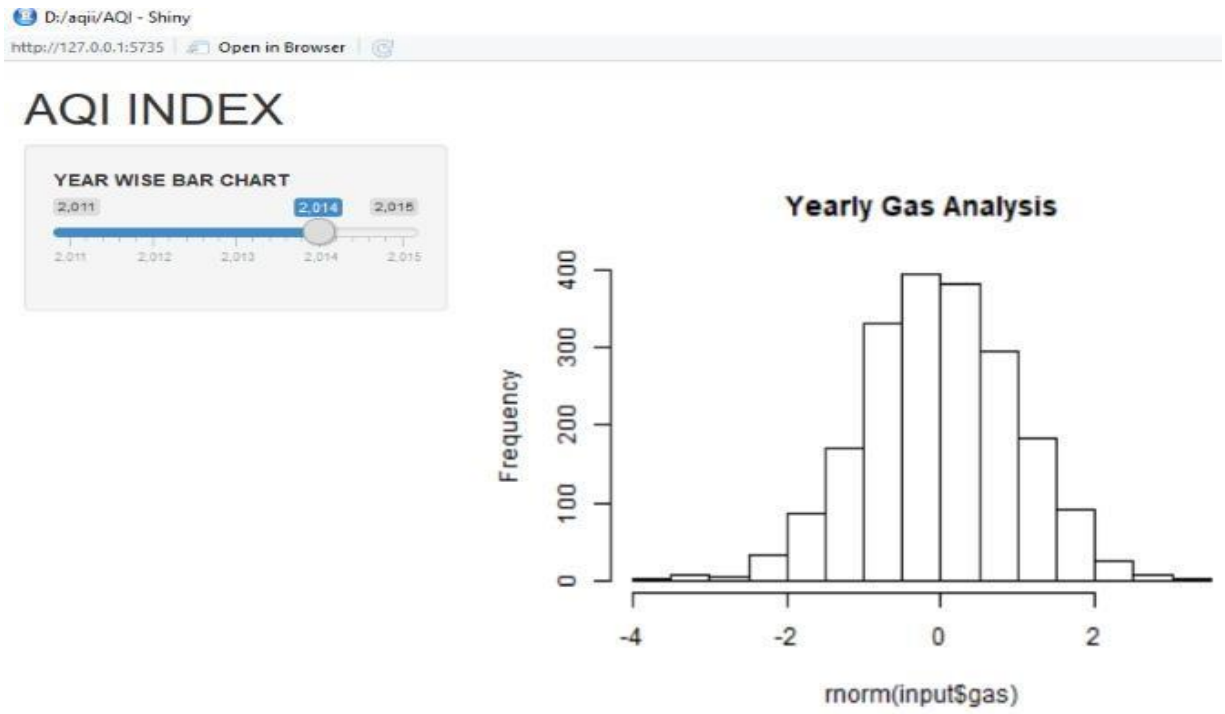

Figure 5. *AQI Index and yearly Gas Analysis (Snap Shot-2)*

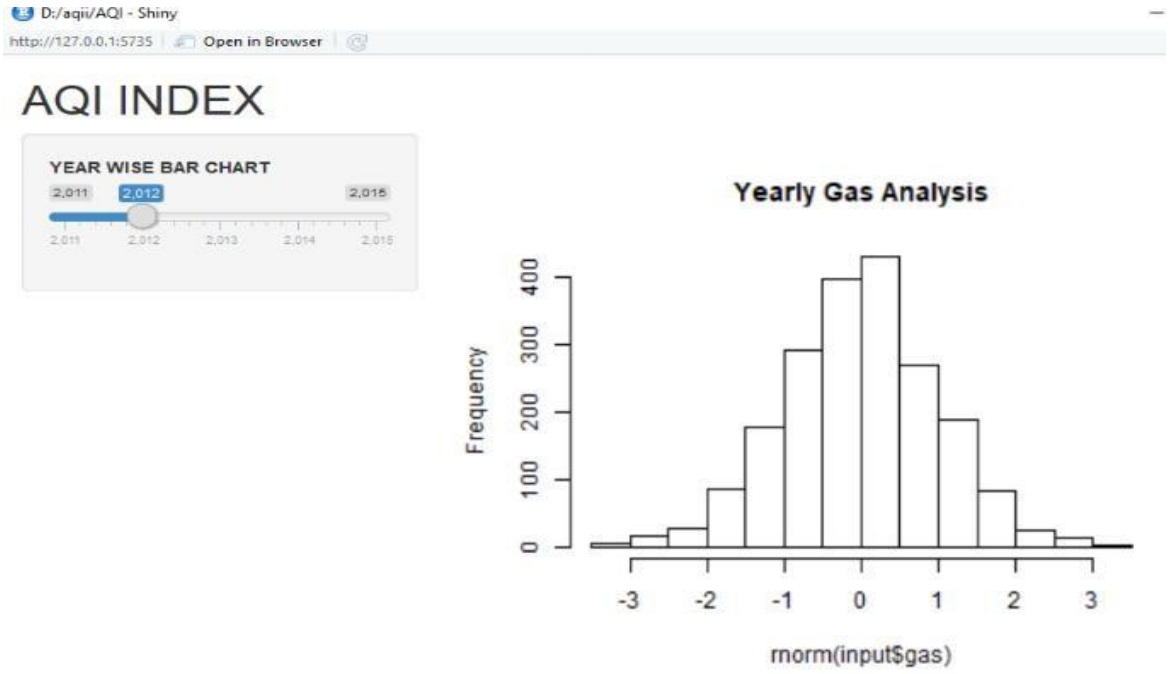

Figure 6. *AQI Index and yearly Gas Analysis (Snap Shot-3)*

6.3 Test Cases

Table 1. Different Test cases of results

Case Study and Setup for the Experiment	Different Scenarios	Points before the Test	Cascades for Exp.	Info Collected from Exp.	Outcome Tentative	Fact and Real Outcome	Ok / Not Ok
1. Checking for calculating accuracy	Check Response on entering valid dataset	R Studio must be pre-installed	Pass the dataset file	Dataset in csv format	Display the accuracy	Display the accuracy	Pass
2. Predict the result	Check response on entering 4 gases.	R Studio must be pre-installed	Run the application	Enter different gases	Good/ moderate/ Poor/ Severe	Good/ moderate/ Poor/ Severe	Pass

Observed Accuracy:

Table 2. Accuracy Analysis in the Experiements

K-means Clustering	80% -90%
Naïve Bayes	80% - 85%

6.4 Results

Pl. refer the result Snapshot Fig. 7 to Fig. 12 for the information and data collected and quantified after implementing this info in the Analytics Algo on Dataset Available.

```
View(X2011)
> attach(X2011)
> summary(SO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   2.000   4.000   4.000   5.325   5.000  33.000       4
> summary(NO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   17.00   43.00   53.00   58.49   66.00  151.00       4
>
> summary(`RSPM/PM10`)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    27.0   127.0   200.0   222.2   296.0   692.0      11
> summary(SPM)
    Length     Class      Mode
       736 character character
```

Figure 7. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-1)*

```
View(X2012)
> attach(X2012)
> summary(SO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   2.000   2.000   4.000   7.444  10.000  49.000
>
> summary(NO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   20.00   59.00   83.00   81.68  102.00  189.00
>
> summary(`RSPM/PM10`)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    22.0   136.8   210.0   236.8   303.2   843.0
> summary(SPM)
    Length     Class      Mode
       288 character character
```

Figure 8. *Data Collected and Quantified after implementing info into Analytics Algo on Dataset Available (Snapshot-2)*

```
View(X2013)
> attach(X2013)
> summary(SO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   2.000   2.000   3.000   4.487   6.000  40.000
> summary(NO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   24.00   70.00   89.00   89.73  111.00  199.00
> summary(`RSPM/PM10`)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    45.0   123.5   223.0   228.2   303.5   670.0
> summary(SPM)
    Length     Class      Mode
       271 character character
```

Figure 9. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-3)*

```
 View(X2014)
> attach(X2014)


> summary(SO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   2.000   4.000   4.000   5.137   4.000  28.000
> summary(NO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   11.00   41.00   54.00   63.49   72.00  209.00
> summary(`RSPM/PM10`)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    21.0   137.0   194.0   217.1   273.0   709.0      13
```

Figure 10. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-4)*

```
View(X2015)
> attach(X2015)


> summary(SO2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  2.000   4.000   4.000   5.639   4.000  46.000      13
> summary(NO2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  18.00   45.00   53.00   67.36   80.00  224.00      12
> summary(`RSPM/PM10`)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   18.0   149.0   210.0   218.8   271.0   892.0      11
```

Figure 11. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-5)*

```
SUMMARY OF 5 YEARS POLLUTION GASES

summary(AQI$SO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   2.000   4.000   4.000   5.498   5.000  49.000      17
> summary(AQI$NO2)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   11.00   45.00   57.00   67.68   85.00  224.00      16
> summary(AQI$PM10)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    18.0   137.0   204.0   222.1   286.0   892.0      35
> summary(AQI$SPM)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    0.00    0.00    0.00   12.75    0.00  318.00
```

Figure 12. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-6)*

6.5 K-means Algorithm Results:

K-means:

The K-means algorithm is based on few steps:

6.5.1 Objects/ entities in no. k from the overall populations which will define the first hand cluster definitions.

6.5.2 The objects are assigned the same cluster, depending on distance each obstacle and cluster center (centroids).

6.5.3 Calculation and recording of new mean of individual cluster.

6.5.4 The process is then iterated till the finalized condition of the criteria is obtained upto acceptable range.

Screenshot of Algorithm Used (Refer Fig. 13 to Fig. 16)

```
>results <- kmeans(na.omit(AQI.features),3)
>results
K-means clustering with 3 clusters of sizes 1224, 395, 1108

Cluster means:
        SO2      NO2     PM10      SPM
1 4.355392 56.96324 123.8243  8.854575
2 7.556962 92.59241 436.0886 10.215190
3 6.056859 71.12004 254.0686 17.648014

Clustering vector:
   [1] 1 3 3 3 3 3 3 3 1 1 3 3 1 1 1 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 1 1 1 2 1 1 1 3 1 1 3 1
  [44] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 3 3 2 1 3 3 3 3 2 3 3 3 3 2 3 2 2 1 3 3 1 3
  [87] 1 1 1 3 1 3 3 1 1 1 1 3 1 3 3 1 3 3 3 3 2 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [130] 3 2 1 3 2 3 3 2 3 3 3 3 3 2 3 2 3 3 2 2 3 2 3 1 1 1 1 1 1 1 1 3 3 2 3 2 3 3 3 1 1 3
 [173] 1 3 1 3 3 3 3 3 3 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 3 3 1 1 3 3 1 3 2 3
 [216] 2 3 2 2 3 3 3 3 3 3 1 3 1 3 1 1 1 1 1 1 1 1 3 3 1 3 3 3 3 3 3 3 3 1 1 1 3 1 1 1 1 1
 [259] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 1 3 2 3 3 3 3 3 1 3 1 2 2 3 3 2 3 3 2 3 2 3 1 1 3
 [302] 3 1 3 3 3 3 3 3 2 2 3 3 3 3 3 3 1 3 1 3 1 1 1 3 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 3
 [345] 3 2 2 2 2 2 3 2 2 2 3 2 2 3 3 2 2 3 2 2 3 2 2 2 3 3 3 3 1 3 3 2 3 3 3 1 2 3 3 3 3 3 3 3
 [388] 3 1 3 3 1 1 1 1 1 3 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 2 2 3 2 2 3 2 3 2 2 3 3 2
 [431] 2 2 2 1 1 3 2 2 1 1 3 3 1 1 3 3 1 1 1 1 1 1 1 1 3 1 3 1 1 1 1 1 3 2 2 3 1 1 1 1 1 3 1
 [474] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 1
 [517] 3 3 3 3 3 3 2 2 3 2 3 2 2 1 1 2 2 2 3 3 2 3 2 2 3 3 3 1 3 3 3 3 1 3 2 3 3 3 3 3 3 2 3
 [560] 3 3 2 2 2 2 2 2 1 2 3 1 1 2 3 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Figure 13. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-1) Used K-Means*

```
[ reached getOption("max.print") -- omitted 1727 entries ]

Within cluster sum of squares by cluster:
1] 3683561 4204609 5108828
 (between_SS / total_SS =  85.6 %)

Available components:

1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss""betweenss"
7] "size"         "iter"         "ifault"
>results$size
1] 1224  395 1108
>results$cluster
   [1] 1 3 3 3 3 3 3 3 1 1 3 3 1 1 1 1 3 3 3 3
  [21] 3 3 1 3 3 3 3 3 3 3 1 1 1 1 2 1 1 1 3 1
  [41] 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
  [61] 2 2 2 2 3 3 2 1 3 3 3 3 2 3 3 3 3 2 3 2
  [81] 2 1 3 3 1 3 1 1 1 3 1 3 3 1 1 1 1 3 1 3
 [101] 3 1 3 3 3 3 2 3 1 1 1 1 1 1 1 1 1 1 1 1
 [121] 1 1 1 1 1 1 1 1 1 3 2 1 3 2 3 3 2 3 3 3
 [141] 3 3 2 3 2 3 3 2 2 3 2 3 1 1 1 1 1 1 1 1
 [161] 1 3 3 2 3 2 3 3 3 1 1 3 1 3 1 3 3 3 3 3
 [181] 3 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1
 [201] 1 1 1 3 3 3 3 1 1 3 3 1 3 2 3 2 3 2 2 3
 [221] 3 3 3 3 3 1 3 1 3 1 1 1 1 1 1 1 1 3 3
 [241] 1 3 3 3 3 3 3 3 3 1 1 1 3 1 1 1 1 1 1 1
 [261] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 1 3 2 3
 [281] 3 3 3 3 1 3 1 2 2 3 3 2 3 3 2 3 2 3 1 1
 [301] 3 3 1 3 3 3 3 3 3 2 2 3 3 3 3 3 3 1 3 1 3
 [321] 1 1 1 3 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1
 [341] 1 1 1 3 3 2 2 2 2 3 2 2 2 3 2 2 3 3 2 2
 [361] 3 2 2 3 2 2 2 3 3 3 3 1 3 3 2 3 3 3 1 2
 [381] 3 3 3 3 3 3 3 3 3 1 3 3 1 1 1 1 1 3 1 3 1
 [401] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 2 2 3
 [421] 2 2 3 2 3 2 2 3 3 2 2 2 2 1 1 3 2 2 1 1
 [441] 3 3 1 1 3 3 1 1 1 1 1 1 1 1 3 1 3 1 1 1
 [461] 1 1 3 2 2 3 1 1 1 1 1 3 1 1 1 1 1 1 1 1
```

Figure 14. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-2) Used K-Means*

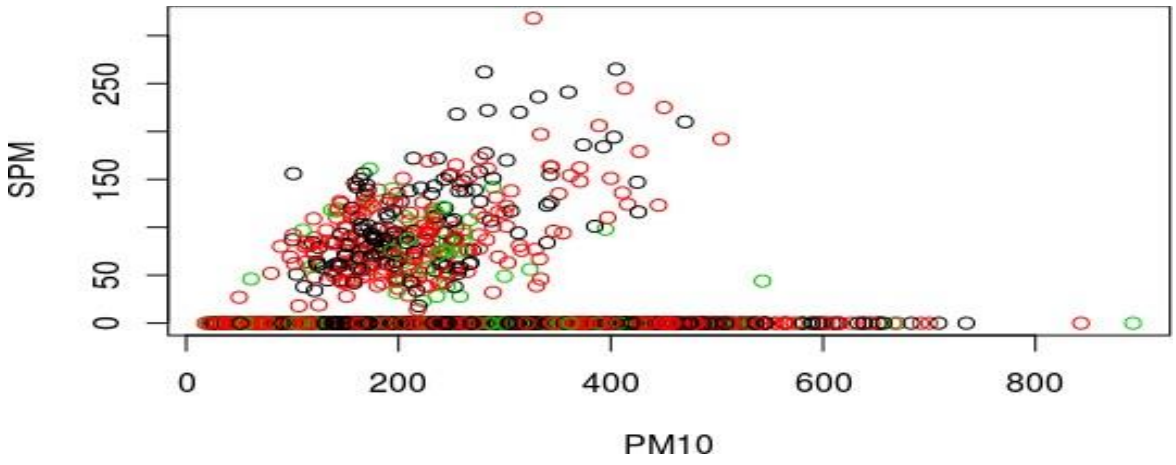

Figure 15. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-3) Used K-Means*

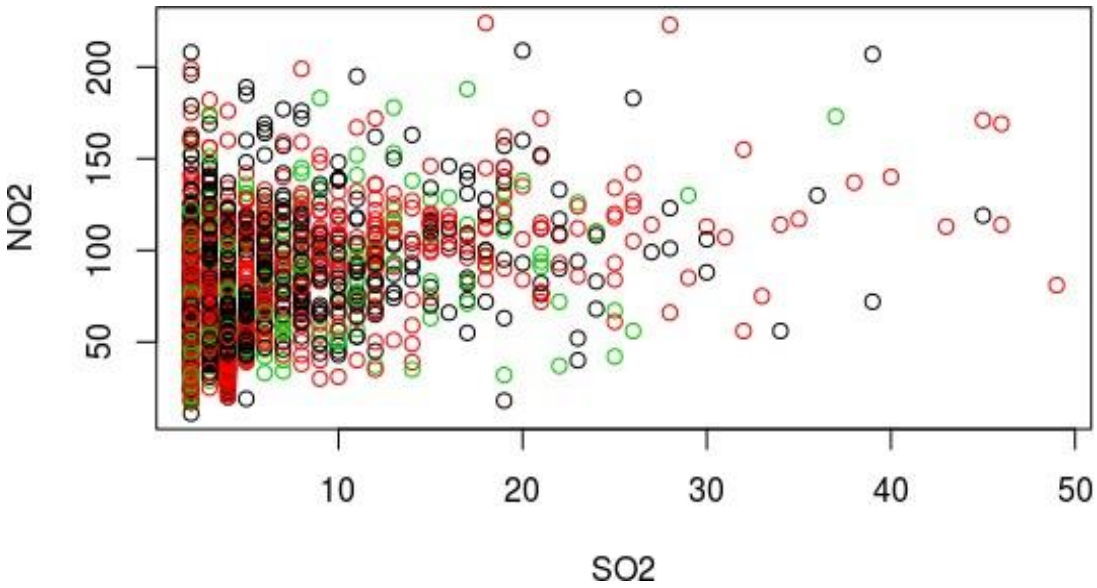

Figure 16. *Data Collected and Quantified after implementing info in the Analytics Algo on Dataset Available (Snapshot-4) Used K-Means*

6.6 Naive Bayes Classification

Naïve Bay's is a popular technique based on probability analysis where the conditional and actual probability of event occurrence is measured. Naïve Bayes defined the probabilistic measurement of happenings and gives very much accurate and realistic outcomes of the possible events. It is based on different initial assumptions and classifies the data into one of the certain categories. (Refer the Fig. 17 to Fig. 19).

Screenshot of Algorithm Used:

```
> pred <- predict(results,vdata)
> pred <- predict(results,vdata)
> head(pred)
          1       15       17       37       40       42
0.10217519 0.10753259 0.10352452 0.10244208 0.10244208 0.09620767
> head(vdata)
    Sampling Date Stn Code State City/Town/Village/Area Location of Monitoring Station
1      01-01-14     531 Delhi             Delhi           Pritampura, Delhi
15     01-01-14     531 Delhi             Delhi           Pritampura, Delhi
17     01-01-15      55 Delhi             Delhi           Nizamuddin, Delhi
37     01-02-14     531 Delhi             Delhi           Pritampura, Delhi
40     01-02-14     531 Delhi             Delhi           Pritampura, Delhi
42     01-02-14     531 Delhi             Delhi           Pritampura, Delhi
                                  Agency             Type of Location SO2 NO2 PM10
SPM
1  Central Pollution Control Board Residential, Rural and other Areas   4  40  154   0
15 Central Pollution Control Board Residential, Rural and other Areas   4  41  232   0
17 Central Pollution Control Board Residential, Rural and other Areas   4  44  203   0
37 Central Pollution Control Board Residential, Rural and other Areas   4  37  134   0
40 Central Pollution Control Board Residential, Rural and other Areas   4  37  134   0
42 Central Pollution Control Board Residential, Rural and other Areas   4  45  115   0
               _conds _dewptm _fog _hail _heatindexm _hum _precipm _pressurem _rain
_snow _tempm
1          Mist     13    0    0      <NA>   98    <NA>      1020     0    0    13
15 Patches of Fog      9    1    0      <NA>   98    <NA>      1021     0    0     9
17         Smoke      11    0    0      <NA>   54    <NA>      1019     0    0    18
37          Mist      11    0    0      <NA>   91    <NA>      1020     0    0    12
40         Smoke      13    0    0      <NA>   64    <NA>      1018     0    0    18
```

Figure 17. *Numerical Results Obtained after Applying the Data Analytics Naïve Baye's Classification Algo on Dataset Available (Snapshot-1)*

```
 AQI_nb<-naiveBayes(nPM10 ~
Sampling.Date+Location.of.Monitoring.Station+Agency+Type.of.Location+X_conds+X_de
wptm+X_fog+X_hail+X_hum+X_pressurem+X_rain+X_tempm
+X_thunder+X_tornado+X_vism+X_wdird+X_wdire+X_wspdm  , data = AQItrain)
>
> AQI_nb

Naive Bayes Classifier for Discrete Predictors

Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:
Y
      High       Low
0.5755475 0.4244525

Conditional probabilities:
        Sampling.Date
Y        01-01-14  01-01-15  01-02-14  01-03-11  01-03-12  01-03-13  01-03-14
01-04-11  01-04-12
        Sampling.Date
Y        01-04-13  01-04-15  01-05-11  01-05-13  01-05-15  01-06-11  01-06-14
01-06-15  01-07-11
        Sampling.Date
Y        01-07-12  01-07-14  01-07-15  01-08-12  01-08-13  01-08-14  01-09-13
01-09-14  01-09-15
        Sampling.Date
Y        01-10-11  01-10-12  01-10-13  01-10-14  01-10-15  01-11-11  01-11-12
01-12-11  01-12-12
```

Figure 18. *Numerical Results Obtained after Applying the Data Analytics Naïve Baye's Classification Algo on Dataset Available (Snapshot-2)*

```
 [ reached getOption("max.print") -- omitted 1727 entries ]

Within cluster sum of squares by cluster:
[1] 3683561 4204609 5108828
 (between_SS / total_SS =  85.6 %)

Available components:

[1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss""betwe
[7] "size"         "iter"         "ifault"
>results$size
[1] 1224  395 1108
>results$cluster
    [1] 1 3 3 3 3 3 3 3 1 1 3 3 1 1 1 1 3 3 3 3
   [21] 3 3 1 3 3 3 3 3 3 3 1 1 1 1 2 1 1 1 3 1
   [41] 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   [61] 2 2 2 2 3 3 2 1 3 3 3 3 2 3 3 3 3 2 3 2
   [81] 2 1 3 3 1 3 1 1 1 3 1 3 3 1 1 1 1 3 1 3
  [101] 3 1 3 3 3 3 2 3 1 1 1 1 1 1 1 1 1 1 1 1
  [121] 1 1 1 1 1 1 1 1 1 1 3 2 1 3 2 3 3 2 3 3 3
  [141] 3 3 2 3 2 3 3 2 2 3 2 3 1 1 1 1 1 1 1 1
  [161] 1 3 3 2 3 2 3 3 3 1 1 3 1 3 1 3 3 3 3 3
  [181] 3 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1
  [201] 1 1 1 3 3 3 3 1 1 3 3 1 3 2 3 2 3 2 2 3
  [221] 3 3 3 3 3 1 3 1 3 1 1 1 1 1 1 1 1 1 3 3
  [241] 1 3 3 3 3 3 3 3 3 3 1 1 1 3 1 1 1 1 1 1
  [261] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 1 3 2 3
  [281] 3 3 3 3 3 1 3 1 2 2 3 3 2 3 3 2 3 2 3 1 1
  [301] 3 3 1 3 3 3 3 3 2 2 3 3 3 3 3 3 1 3 1 3
  [321] 1 1 1 3 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1
  [341] 1 1 1 3 3 2 2 2 2 3 2 2 2 3 2 2 3 3 2 2
  [361] 3 2 2 3 2 2 2 3 3 3 3 1 3 3 2 3 3 3 1 2
  [381] 3 3 3 3 3 3 3 3 3 1 3 3 1 1 1 1 1 3 1 3 1
  [401] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 2 2 3
  [421] 2 2 3 2 3 2 2 2 3 3 2 2 2 2 1 1 3 2 2 1 1
  [441] 3 3 1 1 3 3 1 1 1 1 1 1 1 1 3 1 3 1 1 1
  [461] 1 1 3 2 2 3 1 1 1 1 1 3 1 1 1 1 1 1 1 1
  [481] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Figure 19. *Numerical Results Obtained after Applying the Data Analytics Naïve Baye's Classification Algo on Dataset Available (Snapshot-3)*

7 Recommendations and Suggestions

- If the observed value of the gas is 2 times lesser than the mean then it is healthy to go outside and categorized as GOOD.

Precaution-None.

- If the observed value of the gas is 1 time lesser than the mean value of the calculated gas then it is categorized as MODERATE.

Precaution-Sensitive people should reduce outdoor exertion.

- If the observed value of the gas is around the mean value of the calculated gas then it is categorized as UNHEALTHY.

Precaution- People with diseases and children, adults should reduce outdoor exertion.

- If the observed value of the gas is twice than the mean value of the calculated gas then it is categorized as VERYUNHEALTHY.

Precaution- People with diseases and children, adults should avoid outdoor exertion. Others should limit their exertion.

- If the observed value of the gas is thrice than the mean value of the calculated gas then it is categorized as HAZARDOUS.

Precaution-People with diseases, who are active outside mostly and children, adults should reduce outdoor exertion. Others should limit their exertion.

The Government of National Capital Territory of Delhi has made totally numerous steps for the duration of the cutting-edge decade to decrease the extent of air harms within the city. The upsides of air sullying control estimations are clear inside the readings. The final touch of the website is better and still higher to authorize. Regardless, notably after this it need to be performed to decrease the level of air sullying (Rastogi, R. et al., 2023c).

Likewise better results need to typical. There is a need to reinforce appropriately particular measures and to scale up. Government tries by myself are deficient. Society's collaboration is primary to have an impact inside the decline of poisons. There is a need to propel the use of open vehicle. The use of metro stations may be improved with the aid of the route of motion of a careful quantity of feeder transports at metro-stations operating with needed stations.

Prosperity, as we overall know, is a predictable subject, not simply inside the sets of the Department of Health, anyway with every last one of the people who are busy with human progression. Various unprecedented specialists have commented on the criticalness of nature in the quality of every person. Consequently, whatever people accepts a vocation in changing the earth in any better way, for no good reason, need to add to protecting the prosperity of people by totally controlling all of the components that stress it.

8 Conclusion

It is suggested that for BY and huge and predictable air quality stations, the AQI is held near the most steady for what can be normal for those conditions. Manual stations and for every area, month-to-month representing AQI is done to guarantee that manual data for AQI is checked and got to. An online AQI spread structure is intended for some flashy, clear and coordinated looking AQI questions. The site's different features and properties that incorporate the statement of harmful poisons for the record outperform the toxic substance standards and prosperity impacts (Rastogi, R. et al., 2023b).

There is a need to instruct people to kill their vehicles at the same time as retaining up at traffic crossing focuses. The advancement in ousted individuals' inflows can be faded through making and making business openings in city locales and country areas, in this manner preventing further sticks of the accurately blocked capital Delhi.

References

1. Azam, A.G., Zanjani, B.R., and Mood, M.B. (2016), Effects of air pollution on human health and practical measures for prevention in Iran, *J Res Med Sci.*; 21:65.

2. Rai, P. (2016). Biodiversity of roadside plants and their response to air pollution in an Indo-Burma hotspot region, implications for urban ecosystem restoration, pp. 47-55.

3. Veni, K., Marimuthu, Lavanya, K., (2014). Air Pollution Tolerance Index of Plants: A Comparative Study, *IJPPS (0975-1491), Vol. 6*, issue 5.

4. World Health Organization. (18 June 2010). Anniversary of smallpox eradication. Geneva.

5. A. Masih (Automn 2019). Machine Learning algorithm in air quality modeling, *Global journal of environmental science and management, Volume 5*, Issue 4, Pages 515-534.

6. A. Masih (Summer 2019). Application of ensemble learning techniques to model to atmospheric concentration of SO2, *Global journal of environmental science and management, Volume 5*, Issue 3, Pages 309-318.

7. Rastogi, R., Saxena, M., Gupta, M., Rastogi, A. R., Kumar, P., Jain, M., Rastogi, M., Gupta, C., Tyagi, A., & Srivatava, P. (2021). Happiness Index and Gadget Radiation Analysis on Yajna and Mantra Chanting Therapy in South Asian Continent: COVID-19 vs. Ancient Rich Culture From Vedic Science. *International Journal of Health Systems and Translational Medicine (IJHSTM), 1*(1), 1-46. http://doi.org/10.4018/IJHSTM.2021010101

8. Dash, S., Shakyawar, S.K., Sharma, M. et al (2019).. Big data in healthcare: management, analysis and future prospects. J Big Data 6, 54, doi:10.1186/s40537-019-0217-0

9. Suliankatchi, R., Baridalyne, N. & Gupta, S., (2013). Air pollution in Delhi: Its Magnitude and Effects on Health. Indian Journal of Community Medicine. 38. 4-8.

10. Brondino, N., De Silvestri, A., Re, S., Lanati, N., Thiemann, P., Verna, A., Politi, P. (2013).A Systematic Review and Meta-Analysis of Ginkgo biloba in Neuropsychiatric Disorders: From

Ancient Tradition to Modern-Day Medicine, Evidence-Based Complementary and Alternative Medicine, (1), 1–11. https://doi.org/10.1155/2013/915691\

11. Chauhan, S., Rastogi, R., Chaturvedi, D.K., Arora, N., Trivedi, P. (2017). Framework for Use of Machine Intelligence on Clinical Psychology to study the effects of Spiritual tools on Human Behavior and Psychic Challenges. Proceedings of NSC-2017(National system conference), DEI, Agra, Dec. 1-3, 2017.

12. Sharma, A., Rastogi, R., Chaturvedi, D.K., Satya, S., Arora, N., Trivedi, P., Singh, A., Singh, A. (2019). Intelligent Analysis for Personality Detection on Various Indicators by Clinical Reliable Psychological TTH and Stress Surveys, in the proceedings of CIPR 2019 at Indian Institute of Engineering Science and Technology, Shibpur on 19th-20th January 2019, Springer-AISC Series.

13. Kasthuri, A. (July-Sept.-2018). Challenges to Healthcare in India - The Five A's, Indian Journal of Community Medicine, 141-143, Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6166510/

14. Rastogi, R., Chaturvedi,D.K. (2023a). Mantra and Homa Therapy: Computational Analysis of Different Aspects to Benefit Mankind With Healthcare 4.0 and Industry, *International Journal of Applied Research on Public Health Management (IJARPHM) 8*(1), 1-24.

15. Rastogi, R., Rastogi, A.R., and Gupta, N. (2022). A statistical analysis of air-pollution before COVID-19: experimental study of three years for intelligent environment conducted at North Indian Zone to extract knowledge, *Interdisciplinary Environmental Review*, *Vol. 22,* No. 3-4, pp 274-291.

16. Rastogi, R., Sagar, S., Singh, B., Tandon, N., T. Rajeshwari, Garg, P., Singh, M., Komal, Dhamija, L., Sharma, M. (2023b). Statistical Analysis of Air Quality by Emission of Different Woods: Facing the Threats of Global Pandemic with Healthcare 4.0, *Int. J. of Indian Culture and Business Management (IJICBM), Vol. 27,* issue 1.

17. Rastogi, R., Tandon, N., Sagar, S., Singh, B., T. Rajeshwari, Rastogi, M., Garg, P., Singh, M., Dhamija, L., (2023c). Analytical study of the effect of agnihotra on AQI and its psycho-social impacts: a perspective amidst second wave of pandemic challenges in National Capital Region of Indian subcontinents, *Int. J. of Indian Culture and Business Management (IJICBM), Vol. 27,* issue 1.