### Seminar on Managing Industrial Air Emissions 21<sup>st</sup> Feb, 2018

**Organized by: EMC & SDG Foundation** 

Supported by : Environmental SA & TSI

## AIR QUALITY ANALYTICS DASHBOARD



**Dr. Prasad Modak** 

## Structure of the presentation

- About environmentalanalytics.co
- Air Quality Analytics Dashboard

## What is environmentalanalytics.co?

A platform will feature tools for analysing and visualizing air quality, stack emissions, water quality, wastewater and noise related BIG data.

Many of these tools have been conceived and developed based on Dr. Modak's extensive research and consulting experience.

## Air Quality Analytics Dashboard

This platform will feature tools for analysing and visualizing air quality, many of these tools have been conceived and developed based on Dr. Modak's extensive research and consulting experience.

## The Dashboard

This is how the opening page of the dashboard looks like.

You will see a list of various tabs on the sidebar which when you select will open that page on the right and lets you perform various functions on the dataset accordingly.

#### **Air Quality Analytics**



## List of tools

- Outliers analysis
- Missing data analysis
- Trend analysis
- Location Importance
   Index (LII)
- Meteorological data analysis
- Map Visualizations
- Violation Analysis
- Temporal Variation

🖷 Home
Choose Station <
•
III View Dataset
🖿 Check Data Quality 🛛 <
O Outliers
≌≊ Missing Data <
월월 Missing Data(summary)
≌≊ Time Series
B≊ Contiguous missing data
Mariation
X Correlation

🛃 Trend Analysis
A Violation Analysis <
🖮 Violation
Lad LII
Meteorology <
🔅 Wind Rose
Pollution Rose
Persistent Wind Rose
🖮 Polar Plot
Map Visualization
Air Quality Index
🔟 Split-page
🛛 Help 🗸 <

# Home Page



**Choose CSV** – You can choose a CSV file from your directory, this 'Browse 'button enables you to choose a file from your directory. Make sure the file data is in the specified format.

**Separator** – the field separator character. Values on each line of the file are separated by this character.

**Quote** – The set of quoting characters. Quoting is only considered for columns read as character, which is all of them unless

### EXCEL

When you open the file in **'Excel'** the system automatically allocates different columns to different variables and hence when you open a CSV file in excel you will see something like this.

### NOTEPAD

When you open the file in '**Notepad**' the system is divides the variables on the basis of their **format like number or character or date and separate them using a 'Separator' like comma/Semi colon/Tab.** 

The character variables are stored using 'Quotes', single quotes/double quotes/None, depending on your system.

	🍋 👗 Cut	Calibr	i	- 11 - J		≡ _   ð	×-	Wrap Text	[	General	<b>.</b>	
Pa	└── ि Copy マ Iste マ	ter B I	<u>u</u> -	 ∃ +   <u>∿</u> +	<u></u>	≡ ≡ €	≠ ≠≡ 🖽	Merge & Ce	enter 👻	<b>₽</b> • % •	€.0 .00 .00 →.0	C: Fo
	Clipboard	Es.	Fon	t	G.		Alignment		E.	Numbe	r G	
A	1 • :	$\times \checkmark$	f <sub>x</sub>	date								
	Α	В	С	D	E	F	G	н	Ι	J	К	
1	date	ws	wd	l nox no2		o3	pm10	so2 co		pm2.5	station.name	
2	01-01-1998 00:00	NA	NA	NA	NA	1	29	4.7225	3.372	5 NA	Station A	
3	01-01-1998 01:00	2.16	230	NA	NA	NA	37	NA	NA	NA	Station A	
4	01-01-1998 02:00	2.76	190	NA	NA	3	34	6.83	9.602	5 NA	Station A	
5	01-01-1998 03:00	2.16	170	493	52	3	35	7.6625	10.217	5 NA	Station A	
6	01-01-1998 04:00	2.4	180	468	78	2	34	8.07	8.912	5 NA	Station A	
7	01-01-1998 05:00	3	190	264	42	0	16	5.505	3.052	5 NA	Station A	
8	01-01-1998 06:00	3	140	171	38	0	11	4.23	2.26	5 NA	Station A	
9	01-01-1998 07:00	3	170	195	51	0	12	3.875	1.99	5 NA	Station A	
10	01-01-1998 08:00	3.36	170	137	42	1	12	3.3475	1.45	5 NA	Station A	
11	01-01-1998 09:00	3.96	170	113	39	2	12	2.9225	1.20	5 NA	Station A	
12	01-01-1998 10:00	6.36	180	100	34	7	10	3.06	0.982	5 NA	Station A	
	01-01-1998 11:00	8.16	190	109	38	8	11	1.7475	1.4	3 NA	Station A	
13	01 01 1550 11.00	0.10										

#### sample.station\_newcr - Notepad

<u>File Edit Fo</u>rmat <u>V</u>iew <u>H</u>elp

date,ws,wd,nox,no2,o3,pm10,so2,co,pm2.5,station.name 01-01-1998 00:00,NA,NA,NA,NA,1,29,4.7225,3.3725,NA,Station A 01-01-1998 01:00,2.16,230,NA,NA,NA,37,NA,NA,NA,Station A 01-01-1998 02:00,2.76,190,NA,NA,3,34,6.83,9.6025,NA,Station A 01-01-1998 03:00,2.16,170,493,52,3,35,7.6625,10.2175,NA,Station A 01-01-1998 04:00,2.4,180,468,78,2,34,8.07,8.9125,NA,Station A 01-01-1998 05:00,3,190,264,42,0,16,5.505,3.0525,NA,Station A 01-01-1998 06:00,3,140,171,38,0,11,4.23,2.265,NA,Station A 01-01-1998 07:00,3,170,195,51,0,12,3.875,1.995,NA,Station A 01-01-1998 08:00,3.36,170,137,42,1,12,3.3475,1.455,NA,Station A 01-01-1998 09:00,3.96,170,113,39,2,12,2.9225,1.205,NA,Station A 01-01-1998 10:00,6.36,180,100,34,7,10,3.06,0.9825,NA,Station A 01-01-1998 11:00,8.16,190,109,38,8,11,1.7475,1.43,NA,Station A 01-01-1998 12:00,7.2,180,110,41,9,13,1.375,1.405,NA,Station A



**Date Selection** – The 'date' column format changes, make sure you select the right date format from the option given. If the date format is wrong the application will not process any results. You can either correct the date format in the original file or choose the correct date format.

**Date Range Selection** – The following tab will allow you to select a date range from within the dataset, If you to analyse data within a certain range from your dataset, you might want to use this option.

**Standard Selection** – This shows the standard selected for calculations in the coming pages like 'Violation analysis'. You can have a look at the standard by clicking the view button

# View Dataset

1<sup>st</sup> Quartile : The first quartile (also called the lower quartile) is the number below which lies the 25 % of the bottom data points.

3<sup>rd</sup> Quartile : The third quartile (also called the upper quartile) has 75 % of the data below it.

Standard deviation : Standard Deviation is a measure of the mean distance of values in a data set from their mean.

Variance : It tells how the data is spread around the mean.

**Coefficient of variation** : It is often expressed as a percentage, and is defined as the ratio of the standard deviation to the mean.



A view of the Data summary tab on Dataset page



Data Summary	-
This table shows an overview of basic statistical values of the parameter present in the data frame.	
	a
Column Show/Hide Copy Download	Search:

**Column Show/Hide :** This button is used to select the columns to show or hide in the table. **Copy :** This button is used to copy data to the clipboard.

**Download :** This button is used to download table in different formats, CSV, Excel or PDF.

**Search** – You can search for data relevant terms or values within the table using the search bar.

date	÷	ws 🌲	wd	Å.	nox		no2	₹	o3	÷	pm10	÷	so2	÷	со	÷	pm2.5 🌲
All		All	All		All	(	All		All		All		All		All		All

Value filter tab : These buttons are used to sort the columns.

**Column Sorting :** This button filter data based on column values.

## What are Outliers ?

Statistical data which is extremely different from the others in the same sample.

#### Mathematically :

An outlier is an observation that lies an abnormal distance from other values in a random sample from a population.



## **Outlier calculation**

Outlier for every parameter is individually calculated on the basis of the average and standard deviation calculation of that parameter. Mean value and the standard deviation of each parameter is calculated from the data,

Mean = Sum of parameter values/ Total number of parameter values

SD is calculated by the formula below:

- x = Parameter values
- x = Mean
- n = Total number of parameter values

 $s = \sqrt{\frac{\sum (x - \overline{x})^2}{n - 1}}$ 



**Outlier values are the values which are : less the (Mean - SD) and greater than (Mean + SD)** 



**Show outliers** – This option will show outlier values in a different colour on the graph.

**Select parameter** – You can select the parameter for which you to analyse the data.

#### **Select average by :** This option will tell you choose on how the mean value of a parameter you want to calculate. It will then average out the data on your selected option like '**Day**', '**Month**' or '**Year**'.

**Sigma** – This will define the variation within which you want to see the outliers from the mean value. Higher sigma will give you less outlier values.

**Remove outliers** – You can select whether you want to remove outlier values from your dataset or not.

## Missing Data Analysis

Missing data, or Missing values, occur when no data value is stored for the variable in an observation. Missing data are a common occurrence and can have a significant effect on the conclusions that can be drawn from the data.

In Air Quality Management, missing data analysis tells us the efficiency of the instrument. If there is a missing data point, It means that the data was not recorded for that hour, which directly tells us for the instances in which the instrument was working or not. Time series, this graph shows you the time series analysis of the missing data in your dataset, the 'Red' lines represent the exact missing data point on the given date and time in your dataset. The 'Yellow' lines represent the exact missing data point below the below the 'Contiguous

#### value'.



The purpose of setting **'Contiguous value'** is for clarity: with long time series it is difficult to see where individual missing hours are. Setting it to 4, for example would show where there are at least 4 hours of continuous missing data.





## **Temporal Variation**

In air pollution, the variation of a parameter by time of day and day of week can reveal useful information concerning the likely sources. For example, road vehicle emissions tend to follow very regular patterns both on a daily and weekly basis. By contrast some industrial emissions or parameters from natural sources (e.g. sea salt aerosol) may well have very different patterns. There are four plots generated : day of the week variation, mean hour of day variation and a **combined** hour of day - day of week plot and a monthly **plot.** Also shown on the plots is the 95 % confidence interval in the mean. These uncertainty limits can be helpful when trying to determine whether one candidate source is different from another.



WEEKDAY AND HOURS PLOT

#### DIURNALPLOT



### MONTHLY PLOT



#### WEEKDAY PLOT



## What is Correlation?

Degree and type of relationship between any two or more quantities (variables) in which they vary together over a period; for example, variation in the level of expenditure or savings with variation in the level of income. A positive correlation exists where the high values of one variable are associated with the high values of the other variable(s). A 'negative correlation' means association of high values of one with the low values of the other(s). Correlation can vary from +1 to -1. Values close to +1 indicate a high-degree of positive correlation, and values close to -1 indicate a high degree of negative correlation.

### **Examples of the 'colour' visualization method**

### **Visualization method**

'Number' shows the correlation coefficient values. 'Colour' shows the correlation matrix in shades of different colours with the colour code depending on the extent of correlation

### **Choose Layout**

*'Upper triangle' displays the upper triangular matrix. 'Full' displays the full correlation matrix.* 



### Upper triangle



## **Trend Analysis**

Calculating trends for air parameters is one of the most important and common tasks. Sometimes it is useful to have a general idea about how concentrations might have changed.

Theil-Sen estimate of the slope is the median of all these slopes. The advantage of the using the Theil-Sen estimator is that it tends to yield accurate confidence intervals even with non-normal data and heteroscedasticity (non-constant error variance). It is also resistant to outliers - both characteristics can be important in air pollution.

Scatter plots with conditioning and three main approaches: conventional scatter plot, hexagonal binning and kernel density estimates. It is flexible enough to consider lots of conditioning variables and takes care of fitting smooth or linear relationships to the data.

#### **Examples of trend plots using different methods**

#### **Scatter plot method**

#### **Theil-Sen method**





### Why is Violation Analysis important?

Violated data points are those which lie above the defined average standard values. The standard range is different for every parameter. The standards considered for our calculation are defined by the Central Pollution Control Board (CPCB) standards.

**Simultaneous violations :** A simultaneous violation option is also given to check for simultaneous violations of the selected parameters



#### Location Importance Index (LII)

Location Importance Index or LII is a technique to measure significance of location of monitoring site by comparison of two or more sites on the performance of a

 single parameter at a time and
 two or more parameters simultaneously.
 Missing percentage are only calculated for Contiguous exceedance. Contiguous exceedance is the number of continuous exceeded data points.



A map depicting the location of stations on the LII page from the dashboard.

### How is LII calculated?





An output of comparison of LII values graphically on the dashboard

### WIND ROSE

The wind rose is a very useful way of summarising meteorological data. It is particularly useful for showing how wind speed and wind direction conditions vary by year. Plots are generated for wind roses in a variety of ways: summarising all available wind speed and wind direction data, plotting individual wind roses by year, and also by month. The latter is useful for considering how meteorological conditions vary by season.



Example of Wind Rose plot generated on the dashboard

### **POLLUTION ROSE**

Pollution Rose is a variant of Wind Rose that is useful for considering parameter concentrations by wind direction, or more specifically the percentage time the concentration is in a particular range. Pollution Rose can also usefully be used to show which wind directions dominate the overall concentrations.



Example of Pollution Rose plot generated on the dashboard

### PERSISTENT WIND ROSE

This plot gives information of maximum duration of continuous wind direction. This plot helps to detect flaws in the measuring instrument. For example, if a particular wind bin shows a value of 70 hours for a particular direction, it can signal to faulty instruments as the data looks suspicious.



Example of Persistent Wind Rose plot generated on the dashboard

### **POLAR PLOT**

The Polar Plot is bivariate. It is designed to illuminate the effect of both wind speed and wind direction on the movement of air pollution. This plot shows a continuous series of tones that reflect the mean concentrations by changing hue. The colour and intensity at a given location on the plot reflects the concentration.

Polar Plot illustrates the localized concentration mean based on wind direction and speed (how much is present at a specific regions of the plot based on the wind direction and speed and other factors)



Example of **polar plot** generated on the dashboard

### Map Visualizations

This tab allows to visualises the Wind rose, **Pollution rose and** the Polar plots for all the stations in maps. The plots generated will automatically get fetched and one can control the visibility of the plots in the control panel.



A view of the Map Visualization page on the dashboard



**Plot size** – Will let you choose the size of the plots to be made on the maps. The plots made will be the generated plots on the made in the respective tabs.

**Location coordinates**– You must specify the coordinates of station when analysing multiple stations.

## Air Quality Index

Air Quality Index is a tool for effective communication of air quality status to people in terms, which are easy to understand. It transforms complex air quality data of various pollutants into a single number (index value), nomenclature and colour.

Based on the measured ambient concentrations of a pollutant, sub-index is calculated, which is a linear function of concentration (e.g. the sub-index for PM2.5 will be 51 at concentration 31  $\mu$ g/m3, 100 at concentration 60  $\mu$ g/m3, and 75 at concentration of 45  $\mu$ g/m3).

#### A view of the Air Quality Index plots page on the dashboard



### Split Page

This tab will allow you to see and compare two plot outputs on the same page. The plots that were generated on the subsequent pages earlier are list in the list and given as options to choose from.



#### A view of the Map Visualization page on the dashboard

### Help Page



**Help page** – A list of help pages are given for all the various list of tools in the dashboard. You can directly access the help page of a particular tool under the '**Help Tab**'



#### Your may also give a feedback on how your user experience



This form was created inside of Environmental Management Centre LLP. Report Abuse - Terms of Service -Additional Terms

## Air Quality Analytics Dashboard –STACK EMSSIONS

This is how the opening page of the Stack Emissions dashboard looks like.

The Stack Emissions Dashboard offers the same set of functions as there are for the Ambient Air Quality dashboard.

<ul> <li>Home</li> <li>Choose Stack</li> <li>Choose Stack</li> <li>Choose Stack</li> <li>Choose Stack</li> <li>Choose Coale Coale Coale</li> <li>Choose Coale Coale</li></ul>	<complex-block></complex-block>	PMGA		=	Conceived by : Dr.Presed Modek Developed by : Tausif Faronqui (Lead), Vishwa Vijaysheel, Probal S
Choose Stack      Choose Stack Vew Dataset Vew Dataset Vew Vatastors Compute Correlations Analyze Trends Vew Vatastors Compute Correlations Single Coote Double Quobe Double Quobe Petoback Total data	Image: State of the subscription of	Home		User Input	Sample Dataset -
Vew Dataset         Check Data Quaity         Vew Vatistions         Compute Correlations         Analyze Trends         Votation Analysis         Otable Quaity		Choose Stack	*	Select data 6	NOTE: If you open the CSV in excel, it applies its own format on datetime (this system works with CSV date-time format only. The image below is an illustration of difference between excel and notepad format.)
Vew Wataslam         Check Data Quality         Vew Variations         Compute Correlations         Analyze Trends         Votation Analysis         Split-page         Help         Feedback         Totation Analysis         Compute Correlations         Analyze Trends         Votation Analysis         Split-page         Help         Feedback         Totation Analysis         Compute Correlations         Analyze Trends         Votation Analysis         Compute Correlations         Page         Page         Teach         Teach         Mater Markan         Construct         Split-page         Help         Teacht T	Serial Serial Series		-	Browse No file beliected	annen an
Check Data Quality       Comma         Vew Variations       Comma         Compute Corretations       Tab         Compute Corretations       Data         Analyze Trends       Double Quole         Votation Analysis       Double Quole         Split-page       Double Quole         Peedback       Double Quole         Total Data Quality       The Data Quality         Compute Corretations       Double Quole         Votation Analysis       Double Quole         Split-page       Double Quole         Help       Compute Quole         Compute Corretations       Double Quole         Double Quole       Double Quole         Double Quole       Double Quole         Compute Quole       Double Quole         Double Quole       Double Quole	Separator <ul> <li>Commina</li> <li>Sernicokon</li> <li>Tab</li> </ul> Cuote <ul> <li>None</li> <li>Single Cuote</li> <li>Double Cuote</li> <li>Double Cuote</li> </ul> <ul> <li>V/2018 500</li> <li>A the two the two the transmission of the transwission o</li></ul>	View Dataset			at • 1 + 4
Clebck Collar Cuberny       Comma         View Variations       Semicoson         Compute Correlations       Tab         Analyze Trends       None         Single Quote       Single Quote         Double Quote       Double Quote         Double Quote       Double Quote         Double Cuote       Double Quote         Split-poipe       Tread         Help       Comma         Feedback       Tread         Total data       Semicoson         Split-poipe       Tread         Help       Comma         Feedback       Split-poipe         Help       Tread         Total data       Split-poipe         Help       Split-poipe         Help       Tread         Total data       Split-poipe         Help       Split-poipe         Help       Split-poipe         Help       Tread         Total data       Split-poipe         Help       Split-poipe         Help       Split-poipe         Help       Split-poipe         Help       Split-poipe         Help       Split-poipe         Help       Split-poipe	Oris          • Comma           • Semilocolan           • Intervent           • Intervent         • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent           • Intervent	Charles Dates Country		Separator	
View Variations       • Semicodon       • Tab       • Available of the semicodon	• Semicolon       • Tab         • Tab       • V/2018 900 Na       Na <td>Check Data Quality</td> <td></td> <td>O Comma</td> <td>1 date we wil nos not oit pm20 soit or pm25 station.name</td>	Check Data Quality		O Comma	1 date we wil nos not oit pm20 soit or pm25 station.name
<ul> <li>Total Compute Correlations</li> <li>Analyze Trends</li> <li>Volation Analysis</li> <li>Split-page</li> <li>Feedback</li> <li>Feedback</li> <li>Total Cataa</li> <li>Total Cataa</li> <li>Statian Analyze Trends</li> <li>None</li> <li>Single Quote</li> <li>Double Quote</li> <li>Single Cataa</li> <li>Single Cataa</li></ul>	<ul> <li>Total</li> <li>None</li> <li>Single Quote</li> <li>Double Quote</li> <li>Double Quote</li> <li>Total Case</li> <li>To</li></ul>	View Variations	10	Semicolon	2 1/1/2010 0:00 NA NA NA NA 1 29 4.7225 3.3725 NA Station A 3 1/1/2016 1:00 2.16 220 NA NA NA 127 NA NA NA Station A
Compute Correlations         Analyze Trends         Violation Analysis         Split-page         Heip         Feedback         Freedback         Total data         Total data         Download Sample Data         Download Sample Data	Cuote       None       3 10 10 10 10 10 10 10 10 10 10 10 10 10	TOUT THE SHOWED		<ul> <li>Tab</li> </ul>	4 1/1/2016 2:00 2:76 190 NA NA 3 54 6:83 5:0025 NA Station A
Analyze Trends Violation Analysis Split-page Help Feedback Total data No data Modata Modata Modata Modata	• None           • Single Quote           • Single Augustable           • Single Si	Compute Correlations			5 1/1/2016 1:00 2.16 170 493 52 3 15 3.5625 10.2175 NA Station A
Analyze Trends         Analyze Trends         Acialyze Trends         Bouldle Guole         Soft-Dage	● None       Single Cuote         ● Single Cuote       Double Cuote         ● Duble Cuote       Double Cuote         ● The Cuote       The Cuote         • The Cuote			Quote	7 1/1/2016 5x00 3 1390 264 42 0 16 5.505 1.0525 NA Station A
• Single Quote           • Single Quote           • Double Quote           • Double Quote           • Double Quote             • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote             • Polit-page           • Double Quote              • Polit-page           • Double Quote           • Double	• Single Quote           • Double Quote         • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote           • Double Quote         • Double Quote         • Dou	knalvze Trends		<ul> <li>None</li> </ul>	3 1/1/2016 6:00 3 140 171 38 0 11 4:23 2:265 NA Station A
Andatson Analysis <ul> <li>Double Guote</li> <li>Double Guote</li> <li>If V/2016 hoto</li> <liif 2016="" hot<="" td="" v=""><td>Image: Control of the control of th</td><td></td><td></td><td><ul> <li>Single Quote</li> </ul></td><td>10 1/1/2016 9/90 3 1/0 1375 51 0 12 1.875 1.975 NA Station A 30 1/1/2016 9/90 3.34 170 137 42 1 12 3.3475 1.455 NA Station A</td></liif></ul>	Image: Control of the control of th			<ul> <li>Single Quote</li> </ul>	10 1/1/2016 9/90 3 1/0 1375 51 0 12 1.875 1.975 NA Station A 30 1/1/2016 9/90 3.34 170 137 42 1 12 3.3475 1.455 NA Station A
Split-page Help: reedback Total data Total data Do data	Model       Market	Violation Analysis	1561	Double Quote	11 1/1/2016 9:00 1:96 170 118 39 2 12 2:925 1:305 NA Station A
Split-page         Feedback         Feedback         Total data         No data	Total data         No data				
Help Feedback Total data No data Help Feedback Help Feedback Help Feedback Help Foedback Help Foedback Help Foedback Help Foedback Help H	Total data No data Array of the set of	Solit-page			File Edit Format View Help
Help       I       1-1-2014       1000, Ma, Wo, Ma, Wo, Ma, Yay, Yay, Yay, Yay, Yay, Yay, Yay, Ya	In - 2016       ellio y, ma, Ma, Ma, Say, 1, 27, 4, -222, 3, -3725, Ma, Station A         In - 2016       100, yma, Ma, Ma, Ma, Ma, Ma, Ya, Ma, Ma, Station A         In - 2016       100, 2, 15, 120, 44M, Ma, 17, Ma, Ma, Ma, Station A         In - 2016       100, 2, 15, 120, 44M, Ma, 17, Ma, Ma, Ma, Station A         In - 2016       100, 2, 15, 120, 44M, Ma, 17, Ma, Ma, Ma, Station A         In - 2016       100, 2, 15, 120, 44M, Ma, 17, Ma, Ma, Ma, Station A         In - 2016       100, 2, 15, 120, 442, 42, 145, 5365, 38525, Ma, Station A         In - 2016       100, 2, 15, 120, 140, 420, 92, 145, 5365, 38525, Ma, Station A         In - 2016       100, 110, 11, 420, 12, 305, Ma, Station A         In - 2016       100, 110, 11, 420, 12, 305, Ma, Station A         In - 2016       100, 120, 120, 120, 137, 42, 112, 3, 4375, I, 495, Ma, Station A         In - 2016       100, 130, 120, 137, 42, 112, 3, 4375, I, 405, Ma, Station A         In - 2016       100, 130, 120, 137, 42, 112, 3, 4375, I, 405, Ma, Station A         In - 2016       1100, 1100, 130, 120, 137, 42, 112, 3, 4375, I, 405, Ma, Station A         In - 2016       1100, 1100, 130, 120, 137, 42, 112, 3, 4375, I, 435, Ma, Station A         In - 2016       1100, 1100, 14, 17, 130, 14, 11, 204, 51, 140, 140, 5100, 140, 140, 140, 140, 140, 140, 140,	Second Second		a contraction of the second	date,ws,wd,nos,mo2,o3,pe10,so2,co,pe25,station.name
Feedback       1-1-2016 2-00, 130, 466, 130, 546, 530, 548, 548, 548, 548, 548, 548, 548, 548	1-1-2016       2-09, 190, 094, 043, 294, 045, 39, 6025, 08, 55atian A         1-2-2016       3-09, 2-16, 190, 094, 053, 522, 316, 2-125, MA, 55atian A         1-2-2016       3-09, 2-16, 120, 044, 29, 1, 14, 20, 2-15, 140, 2125, MA, 55atian A         1-2-2016       5-09, 2-16, 120, 044, 29, 1, 14, 20, 2-15, 504, 55atian A         1-2-2016       5-109, 11, 202, 2-305, MA, 55atian A         1-2-2016       5-109, 120, 123, 245, 1205, 345, 51, 200, MA, 55atian A         1-2-2016       5-109, 120, 123, 245, 120, 345, 120, 345, 51, 200, MA, 55atian A         1-2-2016       5-109, 120, 123, 245, 120, 51, 200, MA, 5tatian A         1-2-2016       5-109, 120, 35, 120, 113, 392, 348, 51atian A         1-2-2016       5-109, 120, 350, 120, 313, 24, 122, 3, 357, 1.45, MA, 5tatian A         1-2-2016       5-109, 120, 350, 120, 313, 24, 122, 3, 357, 1.45, MA, 5tatian A         1-2-2016       5-109, 120, 350, 120, 325, 120, 300, 51, 300, 130, 51atian A         1-2-2016       5-109, 120, 120, 350, 120, 350, 50, 50, 50, 50, 50, 50, 50, 50, 50,	Help	Sec		1-1-2010 4100,MA,MA,MA,MA,MA,47,27,4.7215,3.3725,MA,5totion A 1-1-2010 1:00,2.16,230,MA,MA,MA,37,MA,MA,MA,Station A
Feedback Feedba	1-1-2016       4.00(2,2,3), 14(3,007), 14(2,07), 14(3,07), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,007), 14(3,07), 1				1-1-2016 2:00,2.76,190,84,84,3,34,6.83,9.6025,84,Statian A
1-1-2815       5:09, 3, 199, 264, 42, 0; 16, 5:395, 3, 8525, NA, Station A         1-2016       6:09, 3, 140, 173, 38, 0; 114, 23, 2; 295, MA, Station A         1-2017       7:09, 30, 512, 0; 113, 39, 2, 145, 149, 149, 38, 5tation A         1-2018       6:09, 3, 36, 179, 135, 34, 112, 3, 3475, 1, 455, MA, Station A         1-2018       6:09, 3, 56, 179, 113, 39, 2, 145, 145, 149, 5tation A         1-2018       100, 30, 51, 709, 113, 39, 2, 102, 2, 145, 145, 145, 146, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 45, 5tation A         1-2018       11:00, 5, 51, 109, 109, 34, 7, 18, 3, 06, 0, 9325, MA, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 44, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 44, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 44, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 44, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 3, 5tation A         1-2018       11:00, 5, 510, 12, 2025, 14, 510, 510         1-2018       11:00, 5, 510, 12, 2025, 14, 510         1-2018       11:00, 5, 510, 12, 2025, 143, 510         1-2018       11:00, 5, 510, 12, 2025, 143, 510         Download Sample Data	1-1-2016 5:00, 1,199,264,42,0,15,5:365,3:3525,Ma,Statton A         1-1-2016 6:00, 1,190,171,30,01,42,2,2:365,Ma,Statton A         1-2-2018 7:00,3,279,195,51,0,12,3:375,1:395,Ma,Statton A         1-2-2018 1:0:00, 1:0,120,127,42,112,3:375,1:395,Ma,Statton A         1-2-2018 1:0:00, 1:0,120,127,42,112,3:375,1:395,Ma,Statton A         1-2-2018 1:0:00, 1:0,120,127,42,112,3:375,1:395,Ma,Statton A         1-2-2018 1:0:00, 1:0,120,120,120,120,120,120,120,120,120,12	Feedback			1-1-2016 4:00,2.4,180,460,75,2,34,0.07,8.9125,86,5tatiun A
1 - 2016       7 (10)       3, 10)       10, 205       10, 205	1-3-2016 7:00/3170/159:551.395,31.995,31.995,31.995,31.995,31.995,31.953,30.5tation A         1-3-2016 7:00/31724,172,3.3475,1.435,30.5tation A         1-3-2016 7:00/3.96,170,117,42,172,3.3475,1.435,00,5tation A         1-3-2016 9:00/3.96,170,117,42,172,3.3475,1.435,00,5tation A         1-3-2016 9:00/3.96,170,117,42,172,3.3475,1.435,00,5tation A         1-3-2016 9:00/3.96,170,117,42,172,3.3475,1.43,00,5tation A         1-3-2016 9:00/3.96,170,101,00,20,30,00,118,00,00,00,00,00,00,00,00,00,00,00,00,00				1-1-2015 5:00,1,190,264,42,0,16,5,505,3.0525,00,Station A
1-1-2016     5:00, 3:0, 120, 137, 42, 1, 12, 3, 435, 1.455, MA, Station A       1-2-2016     5:00, 3:0, 120, 132, 192, 333, 120, 133, 193, 193, 100, 100       1-2-2016     1:00, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1225, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1275, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 30, 121, 1, 1275, 1.20, MA, Station A       1-2-2016     1:1:00, 6:10, 100, 100, 100, 100, 100, 100, 100,	1-1-2016 8:00, 3, 36, 170, 137, 42, 1, 12, 3, 3475, 1.365, MA, Station A         1-1-2016 9:00, 3, 36, 170, 113, 19, 212, 2, 1923, 1, 205, MA, Station A         1-1-2016 9:00, 3, 36, 170, 113, 19, 212, 2, 1923, 1, 205, MA, Station A         1-1-2016 9:00, 3, 36, 170, 113, 19, 212, 2, 1923, MA, Station A         1-1-2016 9:00, 3, 36, 170, 113, 19, 212, 2, 1923, 1, 205, MA, Station A         1-1-2016 9:00, 3, 36, 170, 113, 19, 212, 2, 1923, 1, 205, MA, Station A         1-1-2016 9:00, 36, 9; 100, 100, 34, 9; 112, 1, 7425, 1, 23, MA, Station A         1-1-2016 9:00, 36, 9; 100, 100, 34, 9; 111, 1, 7425, 1, 23, MA, Station A         Download Sample Data         Download Sample Data         No data				1-1-2016 7:00,3,170,195,51,0,12,3.875,1.995,NA,Station A
1-1-2015     10-2015     10-2016	1-1-2015       11:00, 6:.5, 100, 100, 54, 7, 10, 3.00, 0.0323, MA, Statian A         1-1-2015       11:1-2015         1-1-2015       11:1:00, 8:16, 199, 100, 34, 9, 11, 1, 7425, 1:43, MA, Statian A         Download       Sample Data         No data       Stations         No data       No data				1-1-2016 8:00,3.36,120,137,42,1,12,3.3475,1.455,NA,Station A 1-1-2016 9:00,3.96,120,113,39,2,12,2.9235,1.205,NA,Station A
Total data     Stations     Parameters       No data     No data     No data     No data	Total data     Stations     Parameters       No data     No data     No data				1-1-2015 10:00,6.36,100,100,34,7,10,3.06,0.9825,NA,Station A
Total data  No data  Download Sample Data  Parameters  No data  No data	Total data No data No data Download Sample Data Do				1-1-2016 11:00,8.16,190,109,38,8,11,1.7475,1.43,8A,Station A
Total data Stations Parameters	Total data No data Stations Parameters of No data				Download Sample Data
Total data Stations Parameters	Total data No data Stations Parameters of No data				
Total data Stations Parameters	Total data     Stations     Parameters       No data     No data     No data			New case as the case	
No data No data No data	No data No data No data			Total data	Stations Parameters
No data No data No data	No data No data No data			rotar aata	of an office of the second sec
				No data	No data No data

### Real-Time Air Quality Analytics Dashboard

The Dashboard visualizes real-time Air Quality data showing you the air quality levels from the current date till the last 3 days. The dashboard fetches data from an API (https://docs.openaq.org/) and visual analysis is done.

The dashboard captures data on a real-time basis for various stations of Mumbai and Delhi and for a list of 4 pollutants i.e. pm25, pm10, so2 and o3. Time-Series analysis, Outlier Analysis and Violation Analysis are few exaples of the graphical methods to analyse the data.

### **Time Series**

ASV (Above standard Values) – It tells you the percentage of data points that have violated the standard within the last 3 days.

The green dotted line and the yellow line show the 'Mean' and the 'Standard Value' respectively.

### Air Quality Data



### **Outliers**

Outlier Analysis – It shows you the outlier values for the time series of the captured data. The outliers are recalculated for each new data point fetched from the server.



### **Percentage of Outliers**

Sigma value can be varied to check for outliers and the percentage is calculated on the number of points that are outliers with the total number of points.

More the percentage of outliers more fluctuating the air quality levels will be.

### Air Quality Data



### **Boxplots**

This page shows you the boxplot variation for the selected pollutant for each day.

A more comprehensive statistical analysis of the varying pollutant concentrations is done.



### Violations

The violation graph calculates an average value of the particular pollutant for a particular day, 'Height of the bar' represents the value, compares it with the standard (Blue line) and tells you whether it is violated or not. 'Red' means violated, 'Green' means complaint.

#### Air Quality Data



### Maps

The dashboard also captures the location coordinates of the selected station and automatically plots it on a map.

There is also an option of viewing the other graphs with the maps for further analysis.



### **Real-Time Traffic**

You can also see the Real-time traffic around the location of the station. Instant traffic is shown at a particular point of time.



# THANK YOU