Air quality analysis in Tamilnadu

# Introduction:

Poor air quality has severe health implications, leading to respiratory diseases, cardiovascular issues, and other health problems. Air pollution not only affects human health but also harms the environment. It can lead to acid rain, damage to ecosystems, and the deterioration of historical monuments. Tamil Nadu has taken steps to monitor and control air pollution. Government agencies such as the Tamil Nadu Pollution Control Board (TNPCB) oversee air quality regulations and emission standards. Continuous monitoring stations are set up across the state to collect air quality data. Air quality analysis relies on advanced technologies, including air quality monitoring stations, satellite imagery, and data analytics. Innovations in sensor technology and data integration are helping to provide real-time air quality information.

# About phase 4:

In this part we will continue building the project in ways like:

* Perform the air quality analysis and create visualizations.
* Calculate average SO2, NO2, and RSPM/PM10 levels across different monitoring stations, cities, or areas.
* Identify pollution trends and areas with high pollution levels.
* Create visualizations using data visualization libraries (e.g., Matplotlib, Seaborn).

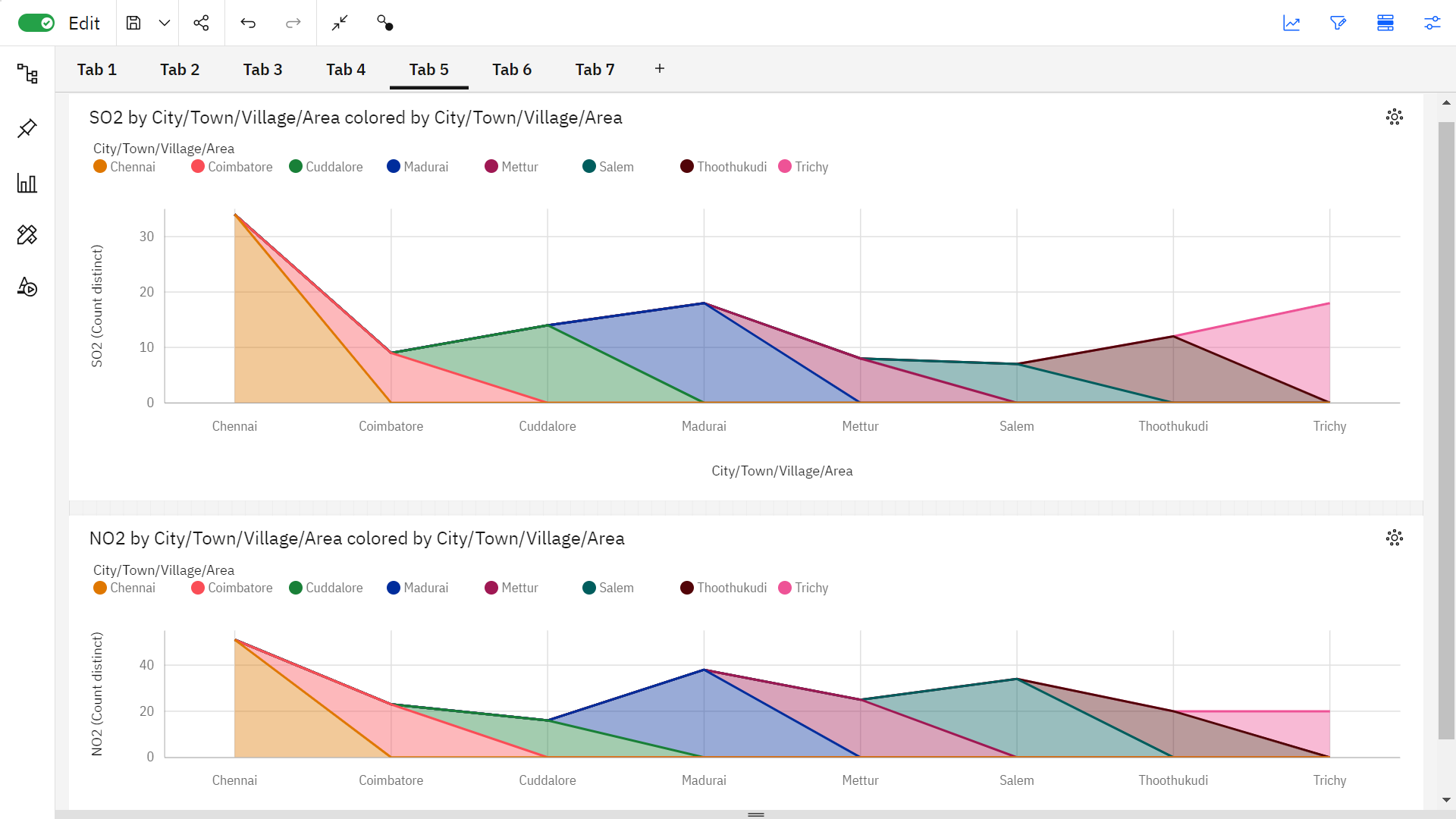
# About IBM cognos:

Cognos is a business intelligence performance management tools for IBM that allows technical and non-technical employees in any company to analyse, extract and create interactive dashboards that enable the company to take relevant key decisions. The Cognos is an intelligence-gathering platform for business that provides an analytical solution for business needs that is scalable and self-service. The highly interactive nature makes it a good way of creating user-friendly dashboards and reports for every company.

# Visualization using IBM cognos:

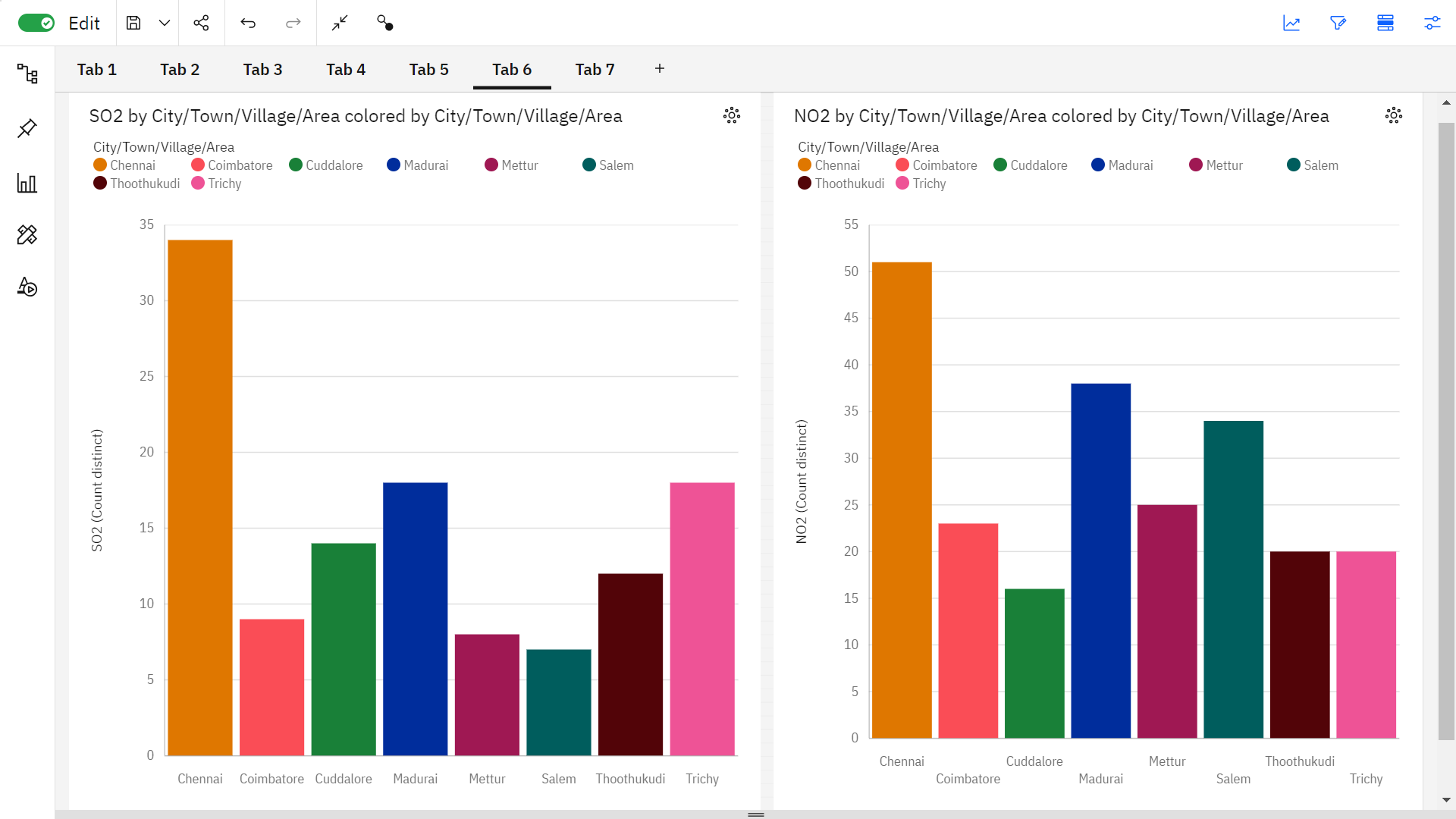
## Area chart:

## An area chart or area graph is based on the line chart but is used to primarily communicate the summation of data rather than represent individual data values, as in line charts. The area between axis and line is usually emphasized with colors, textures or hatchings. The area underneath the line can help one to depict how data progressed with time and can be an excellent way to compare values without going too deep.

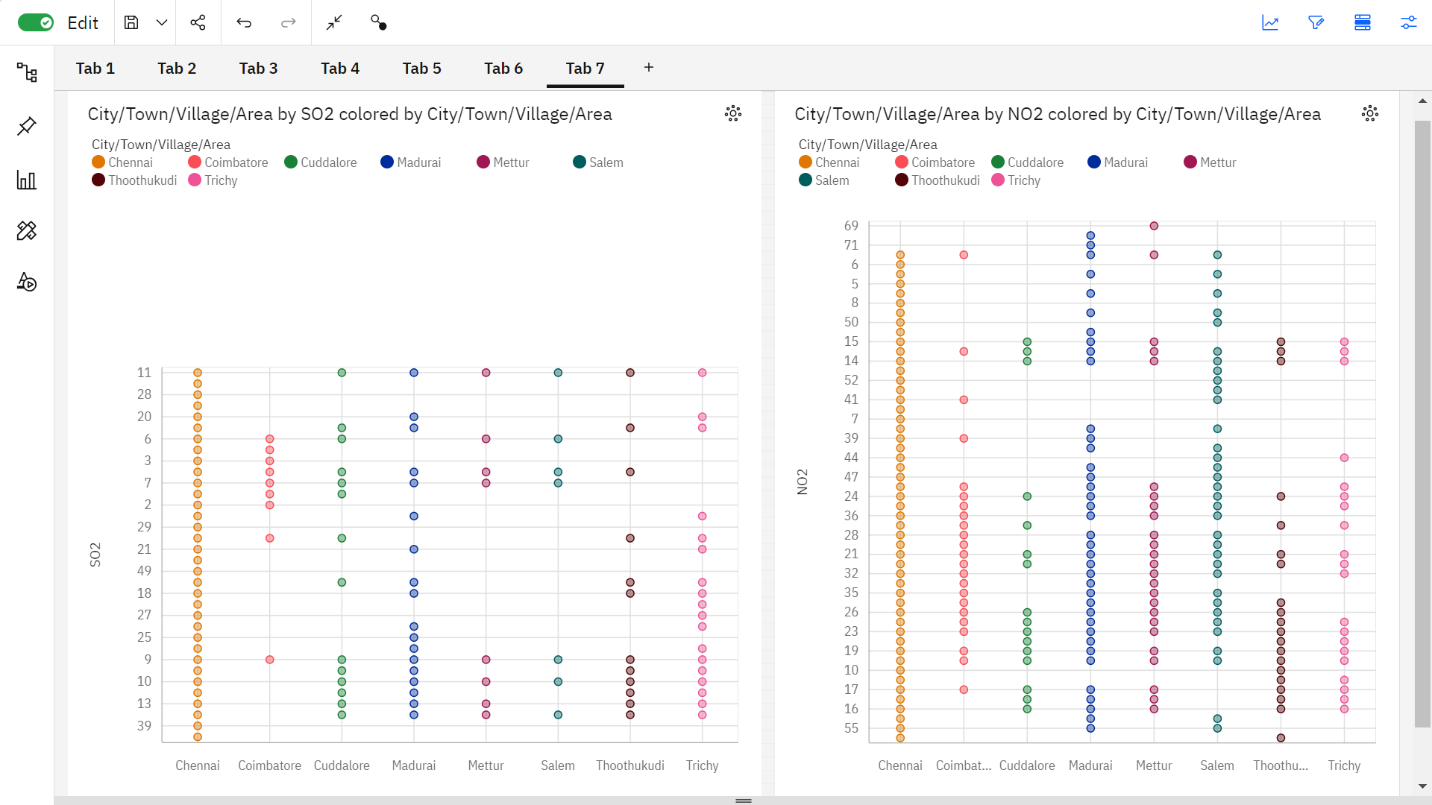


## Stacked column chart:

A stacked column chart is a basic Excel chart type to allow part-to-whole comparisons over time, or across categories. In a stacked column chart, data series are stacked one on top of the other in vertical columns.

Stacked column charts can show change over time because it's easy to compare total column lengths. 

## Bubble chart:

A bubble chart (aka bubble plot) is an extension of the scatter plot used to look at relationships between three numeric variables. Each dot in a bubble chart corresponds with a single data point, and the variables’ values for each point are indicated by horizontal position, vertical position, and dot size. 

## Performance metrics:

Performance metrics are measurable data used to track processes within a business using activities, employee behavior and productivity as key metrics. These metrics track and measure the achievement of overall business goals.

Accuracy is a metric that generally describes how the model performs across all classes. It is useful when all classes are of equal importance. It is calculated as the ratio between the number of correct predictions to the total number of predictions.

The Test accuracy in the above is found to be 1.2.