AIR QUALITY ANALYSIS IN TAMILNADU

DATA ANALYSIS WITH COGNOS

TEAM LEADER- S. YOGESHWARI

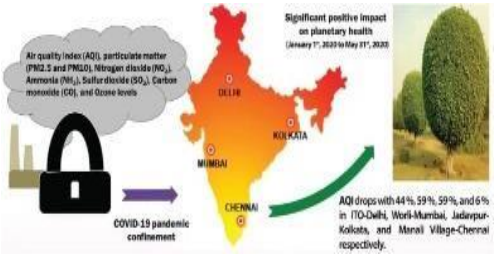
TEAM MEMBER-M. AARTHI

TEAM MEMBER-M. ABINAYA

TEAM MEMBER-D. ABINAYA

PHASE-1

Problem statement and Design thinking



"Air quality analysis refers to the process of monitoring and assessing the composition of the air in a specific area or region to determine the concentration of various pollutants and contaminants present. This analysis is conducted to evaluate the overall quality of the air, identify potential health hazards, and assess compliance with environmental regulations. Air quality analysis typically involves measuring pollutants such as particulate matter, gases (e.g., nitrogen dioxide, sulphur dioxide), volatile organic compounds, and more, with the goal of understanding the impact on human health and the environment."

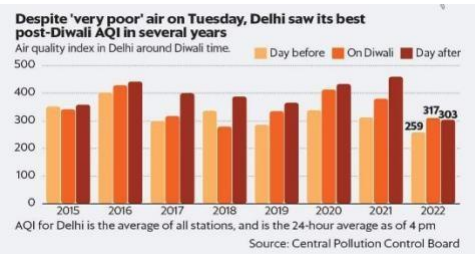
Problem statement:

A problem statement is a concise description of an issue or challenge that needs to be addressed. It typically includes details about the current situation, the specific problem or obstacle, its significance or impact, and sometimes hints at possible solutions. Problem statements are often used in research, project management, and problem-solving processes to clearly define the problem that needs to be solved. They serve as a foundation for devising strategies and solutions.

For example, a problem statement might look like this:

"Many residents in City X are experiencing health issues due to poor air quality caused by high levels of industrial emissions and vehicular pollution. This problem is significant because it leads to a higher incidence of respiratory diseases and decreases the overall quality of life in the community. The problem statement aims to find effective strategies to reduce air pollution and improve public health in City X."

The problem statement sets the stage for further investigation and action to address the identified problem.



Design thinking:

Design thinking is a problem-solving approach and mindset that places a strong emphasis on understanding the needs and perspectives of the end users or customers. It is a creative and iterative process that is often used in fields like product design, innovation, and problem-solving. Here are the key principles and stages of design thinking:

1. Empathize: In this stage, you seek to understand the problem from the user's perspective. This involves empathizing with the end users, observing their behaviour and conducting interviews to gain insights into their needs, motivations, and pain points.

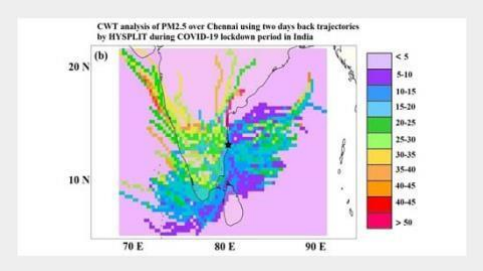
2. Define: Once you have gathered insights, you define the problem statement in a way that focuses on the user's needs. This step involves synthesizing the information you've gathered and identifying the core issues to address.

3. Ideate: This stage encourages brainstorming and generating creative ideas for solving the defined problem. It's about thinking outside the box, encouraging innovation, and considering a wide range of potential solutions.

4. Prototype: Prototyping involves creating tangible representations of your ideas. These can be low models, sketches, or even interactive prototypes, depending on the context. The goal is to quickly test and iterate on your ideas.

5. Test: Testing involves gathering feedback from users by putting your prototypes in front of them. This step helps you assess how well your solutions meet user needs and identify areas for improvement.

6. Iterate: Based on the feedback and insights gained from testing, you refine your ideas and prototypes. This process of iteration may involve going back to previous stages to make necessary adjustments. Design thinking is a flexible approach that encourages collaboration and a user-centric mindset.



PHASE-2

INNOVATION

Step-1. MOBILE AIR QUALITY MONITORING:

https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014

Step-2. AIR QUALITY PREDICTION MODELS:

Create advanced machine learning models that can predict air quality levels in different regions of Tamil Nadu based on historical data, meteorological conditions, and other relevant factors. These predictive models can help residents plan activities and authorities implement timely pollution control measures.

1. Data Collection:

- Gather historical air quality data from various monitoring stations across Tamil Nadu. Ensure that the data includes relevant attributes such as particulate matter (PM2.5 and PM10), nitrogen dioxide (NO2), sulphur dioxide (SO2), ozone (O3), carbon monoxide (CO), and air quality index (AQI).

- Collect meteorological data, including temperature, humidity, wind speed, wind direction, and precipitation, as these factors significantly impact air quality.

2. Data Processing:

- Clean and pre-process the data by handling missing values, outliers, and ensuring data consistency. Normalize or scale numerical features to ensure that they have the same impact on the models.

- Feature engineering: Create new features or derive additional data from existing features. For example, calculate rolling averages, hourly averages, or seasonal effects.

3. Data Splitting:

- Split the dataset into training, validation, and testing sets. The training set is used to train the models, the validation set helps tune hyperparameters, and the testing set is used for final evaluation.

4. Selecting Machine Learning Algorithms:

- Choose appropriate machine learning algorithms for air quality prediction. Time series forecasting models such as ARIMA, Prophet, or machine learning algorithms like Random Forest, Gradient Boosting, Support Vector Regression, or Long Short-Term Memory (LSTM) neural networks are commonly used for this task.

5. Model Training:

- Train multiple machine learning models using the training dataset. Experiment with different algorithms and hyperparameters to find the best-performing model.

6. Model Evaluation:

- Evaluate the models using the validation dataset. Common regression metrics for evaluation include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared.

7. Hyperparameter Tuning:

- Perform hyperparameter tuning to optimize the models. Techniques like grid search or random search can help find the best combination of hyperparameters.

8. Ensemble Methods:

- Consider using ensemble methods like stacking or bagging to combine the predictions of multiple models for improved accuracy.

9. Cross-Validation:

- Apply cross-validation techniques (e.g., k-fold cross-validation) to assess the model's generalization performance and detect overfitting.

10. Real-time Data Integration:

- Once you have a well-performing model, integrate it with real-time meteorological data sources. This enables the model to provide continuous and up-to-date air quality predictions.

11. Model Interpretability:

- Make efforts to interpret and explain model predictions. Techniques like feature importance analysis or SHAP (Shapley Additive explanations) values can help understand the factors driving air quality predictions.

12. Deployment:

- Deploy the trained machine learning model in a production environment, such as a web application or API, so that users can access real-time air quality predictions.

13.Monitoring and Maintenance:

- Continuously monitor the model's performance and retrain it periodically with new data to ensure its accuracy remains high over time.

14. User-Friendly Interface:

- Develop a user-friendly interface (e.g., a web or mobile app) that allows residents and authorities to access air quality predictions and make informed decisions.

15. Public Awareness and Communication:

- Communicate the availability and importance of air quality predictions to the public and relevant authorities. Ensure that the predictions are accessible to a wide audience.

Remember that building accurate air quality prediction models is an ongoing process that may require continuous refinement and adaptation to changing environmental conditions. Collaboration with domain experts and stakeholders can be invaluable in developing effective predictive models for air quality analysis in Tamil Nadu.

Step-3. LOW COST-SENSOR

Establish a low-cost sensor network across urban and rural areas of Tamil Nadu. These sensors can be affordable and easy to deploy, providing real-time data that can be used for analysis and decision-making.

1. Define Objectives and Use Cases:

- Determine the specific objectives and use cases for the sensor network. What type of data are you looking to collect? Examples include air quality, temperature, humidity, soil moisture, water quality, and more.

2. Select Sensor Types:

- Choose appropriate low-cost sensors for each data parameter you want to measure. Ensure that these sensors are affordable, reliable, and easy to deploy.

3. Network Infrastructure:

- Establish a communication infrastructure to collect and transmit data from sensors to a central database. Options include Wi-Fi, Loriann, cellular networks, or a combination depending on the location and coverage.

4. Sensor Deployment:

- Deploy **sensors strategically across urban and rural areas. Consider factors such as population density, environmental conditions, and the purpose of data collection.**

5. Data Aggregation and Storage:

- Set up a central database or cloud-based platform to aggregate and store the data collected from sensors.

6. Data Analysis and Visualization:

- Develop data analysis tools and visualization dashboards to make sense of the data. This can involve creating graphs, maps, and alerts based on sensor readings.

7. Power Supply:

- Ensure a reliable power supply for sensors, especially in rural areas where access to electricity may be limited. Solar panels or battery backup systems can be used.

8. Data Security and Privacy:

- Implement robust security measures to protect the data collected by the sensor network, especially if it includes sensitive information.

9. Maintenance and Calibration:

- Regularly maintain and calibrate sensors to ensure data accuracy and reliability.

10. Community Engagement:

- Engage with local communities, government agencies, and relevant stakeholders to share the benefits of the sensor network and gather input on data usage and prioritises

11. Data Sharing and Accessibility:

Make the collected data accessible to researchers, policymakers, and the public through open data platforms if appropriate.

12. Funding and Sustainability:

- Secure funding for the initial setup and ongoing maintenance of the sensor network. Consider partnerships with government agencies, NGOs, or private sector organizations.

13. Legal and Regulatory Compliance:

- Ensure compliance with any local or national regulations related to data collection and privacy.

14. Continuous Improvement:

- Continuously evaluate the performance of the sensor network and make improvements based on feedback and changing needs.

15. Education and Training:

- Provide training to local technicians and users to ensure the proper operation and utilization of the sensor network.

Establishing a low-cost sensor network is a complex project that requires careful planning, technical expertise, and collaboration with various stakeholders. However, it can provide valuable data for decision-making and contribute to the well-being of urban and rural communities in Tamil Nadu.

Step-4. AIR QUALITY HEALTH INDEX:

Develop a localized Air Quality Health Index (AQHI) that not only measures air quality but also translates it into health impacts specific to Tamil Nadu's population. This can provide more meaningful information to the public.

1. Understand Local Factors:

- Begin by understanding the specific air quality challenges and health risks faced by the population of Tamil Nadu. Consider factors such as industrial pollution, vehicular emissions, agricultural practices, and geographical features.

2. Select Relevant Air Quality Parameters:

- Identify the key air pollutants that have the most significant impact on public health in Tamil Nadu. Common pollutants include PM2.5 (particulate matter), PM10, nitrogen dioxide (NO2), sulfur dioxide (SO2), ozone (O3), and carbon monoxide (CO).

3. Gather Air Quality Data:

- Establish a network of air quality monitoring stations across Tamil Nadu to collect real-time data on these pollutants. Ensure that the data is regularly updated and accurate.

4. Health Impact Assessment:

- Collaborate with health experts and epidemiologists to assess the health impacts associated with exposure to specific pollutant levels. Consider factors like respiratory problems, cardiovascular diseases, and other relevant health issues.

5. Develop Health Risk Categories:

- Create a set of health risk categories that translate air quality data into understandable and actionable information. These categories could range from "Low Risk" to "Hazardous" and include specific health warnings for each category.

6. Public Awareness and Education:

- Educate the public about the AQHI and how to interpret it. Provide information on the health risks associated with different AQHI levels and offer practical advice on reducing exposure during high-risk periods.

7. Online and Mobile Apps:

- Develop a user-friendly website and mobile app that provides real-time AQHI updates and health recommendations based on the current air quality. Make sure the platform is accessible to a wide range of users.

Developing a localized AQHI tailored to Tamil Nadu's specific context can empower the public with information to make informed decisions about their outdoor activities and take precautions during periods of poor air quality. It can also serve as a valuable tool for policymakers and healthcare professionals to address air quality-related health issues in the region.

Step-5. COMMUNITY ENGAGEMENT AND EDUCATION:

Implement an educational program to raise awareness about air quality issues and engage communities in monitoring and improving air quality. This can include school programs, workshops, and citizen science initiatives.

1. Needs Assessment:

- Conduct a needs assessment to understand the specific air quality challenges faced by different communities in Tamil Nadu. Identify areas with high pollution levels and vulnerable populations.

2. Define Educational Goals:

- Clearly define the objectives of the educational program, including raising awareness about air quality, explaining its health impacts, and involving communities in monitoring and mitigation efforts.

3. Curriculum Development:

- Create educational materials, lesson plans, and resources suitable for different age groups and community settings. Develop content that explains the science of air pollution, its sources, and its effects on health and the environment.

4. School Programs:

- Collaborate with schools to integrate air quality education into the curriculum. Offer workshops, presentations, and hands-on activities for students to learn about air quality monitoring and environmental conservation.

5. Teacher Training:

- Train teachers to effectively deliver air quality education materials and activities. Provide them with resources and tools to engage students in meaningful learning experiences.

6. Community Workshops:

- Organize workshops and seminars for community members, including parents, local leaders, and residents, to educate them about air quality issues and solutions. These workshops should be interactive and informative.

7. Citizen Science Initiatives:

- Promote citizen science initiatives where community members can actively participate in air quality monitoring. Distribute low-cost air quality sensors or encourage citizens to use smartphone apps for data collection.

8. Data Analysis and Reporting:

- Teach participants how to analyze air quality data and interpret the results. Create platforms or partnerships where citizens can report their findings and contribute to a larger database.

9. Environmental Clubs and Committees:

- Establish or support environmental clubs or committees within communities, schools, and local organizations. These groups can work together on air quality improvement projects and awareness campaigns.

10. Awareness Campaigns:

- Conduct public awareness campaigns through various media channels, including social media, local newspapers, and community events. Use compelling stories and real-world examples to convey the importance of air quality.

11. Partnerships and Collaboration:

- Collaborate with local environmental organizations, government agencies, and NGOs to strengthen the impact of the educational program. Leverage their expertise and resources for wider reach and sustainability.

12. Evaluation and Feedback:

- Continuously assess the effectiveness of the educational program through surveys, feedback forms, and monitoring air quality improvements. Use this information to refine and improve the program.

13. Policy Advocacy:

- Encourage communities to advocate for policies and regulations that address air quality issues at the local and regional levels. Involve community leaders in these efforts.

By implementing this educational program, you can empower communities in Tamil Nadu to take an active role in monitoring and improving air quality, which can lead to a healthier and more environmentally conscious population.

Step-6. AIR QUALITY MONITORING DRONES:

Use drones equipped with air quality sensors to collect data in hard-to-reach or remote areas. Drones can provide a bird's-eye view of pollution sources and trends.

Using drones equipped with air quality sensors is a innovative and effective way to collect air quality data in hard-to-reach or remote areas, providing valuable insights into pollution sources and trends. Here's a plan for implementing a drone-based air quality monitoring program:

1. Identify Target Areas:

- Determine the specific regions or areas in Tamil Nadu that are hard to access or have limited air quality monitoring infrastructure. Consider factors such as industrial zones, agricultural areas, and areas with vulnerable populations.

2. Select Drone Technology:

- Choose drones that are equipped with appropriate air quality sensors capable of measuring key pollutants such as PM2.5, PM10, NO2, SO2, CO, O3, and VOCs (volatile organic compounds). Ensure that the sensors are calibrated and accurate.

3. Regulatory Compliance:

- Comply with local regulations and obtain the necessary permits for operating drones in the selected areas. Work closely with relevant authorities to ensure legal and safe operations.

4. Flight Planning:

- Develop detailed flight plans that cover the target areas comprehensively. Plan routes that provide optimal coverage and take into account wind patterns, altitude, and safety considerations.

5. Sensor Deployment and Calibration:

- Install air quality sensors on the drones and ensure they are calibrated properly before each flight. Regularly maintain and calibrate sensors to maintain data accuracy.

6.Data collection:

- Using drones equipped with air quality sensors is a innovative and effective way to collect air quality data in hard-to-reach or remote areas, providing valuable insights into pollution sources and trends. Here's a plan for implementing a drone-based air quality monitoring program:

7. Real-Time Data Transmission:

- Implement real-time data transmission capabilities to send air quality data from the drones to a central database or cloud platform. This allows for immediate data analysis and visualization.

8. Data Analysis:

- Analyses the collected data to identify pollution sources, trends, and patterns. Use GIS (Geographic Information System) software to create maps and visualizations that provide a clear understanding of air quality in the target areas.

9. Public Awareness:

- Share the findings and insights from drone-based air quality monitoring with the public through reports, community meetings, and educational materials. Raise awareness about air quality issues and potential health impacts.

10. Policy Recommendations:

- Use the data collected by drones to make policy recommendations to relevant authorities and stakeholders. Advocate for measures to mitigate air pollution in the identified problem areas.

By using drones equipped with air quality sensors, you can gather critical data in areas that were previously difficult to access. This data can contribute to better understanding and management of air quality issues in Tamil Nadu and support efforts to reduce pollution and improve public health.

Step-8. INTEGRATION WITH URBAN PLANNINGS:

Collaborate with urban planners and policymakers to integrate air quality data into city planning processes. This can include zoning regulations, green space development, and traffic management strategies.

1.Engage Stakeholders:

- Identify key stakeholders, including urban planners, city officials, environmental agencies, and community organizations, who can contribute to the integration of air quality data into city planning.

2. Data Sharing and Accessibility:

- Establish a mechanism for sharing real-time air quality data with urban planners and policymakers. Ensure that data is easily accessible through a centralized platform or database.

3. Data Analysis and Visualization:

- Provide tools and resources for data analysis and visualization, such as GIS software, that allow urban planners to explore air quality trends, pollution sources, and hotspots in the city.

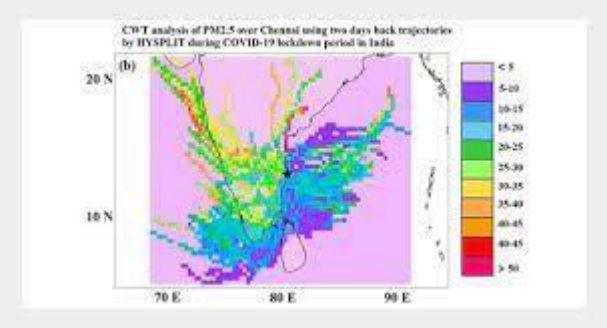
4. Awareness Building:

- Conduct workshops, seminars, and training sessions for urban planners and policymakers to educate them about the importance of air quality in city planning. Highlight the health and economic benefits of clean air.

5. Integration into City Plans:

- Work with urban planners to integrate air quality considerations into existing city plans and policies. This may include updating zoning regulations, land use plans, and transportation strategies to prioritize air quality.

Remember that innovation often involves a combination of technology, community engagement, and policy changes. Collaborate with experts, stakeholders, and local communities to ensure the success and relevance of your project in addressing air quality issues in Tamil Nadu.



.

PHASE-3

DEVELOPMENT PART -1

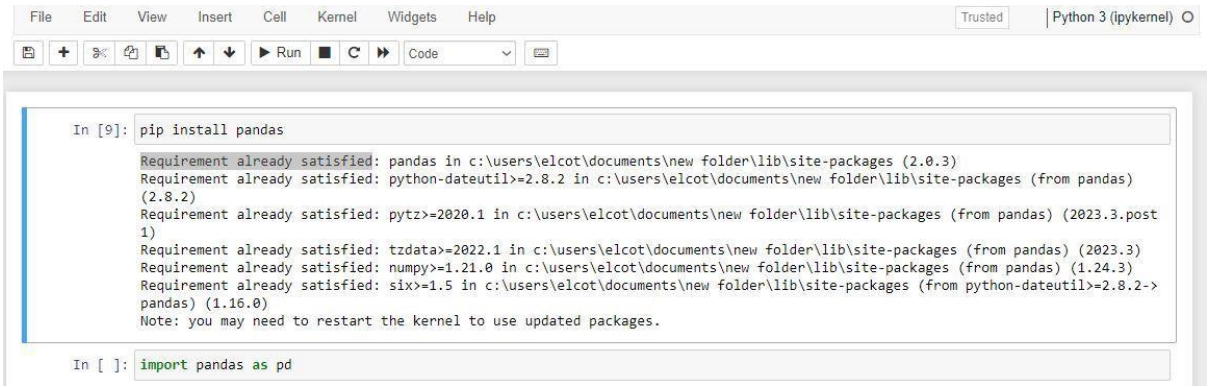
1.Import Libraries:

- First, make sure you have Python and pandas installed. You can install pandas using pip if you don't have it already

```bash
pip install pandas
```

```python
import pandas as pd
```

- A module can be imported into an interactive console environment .



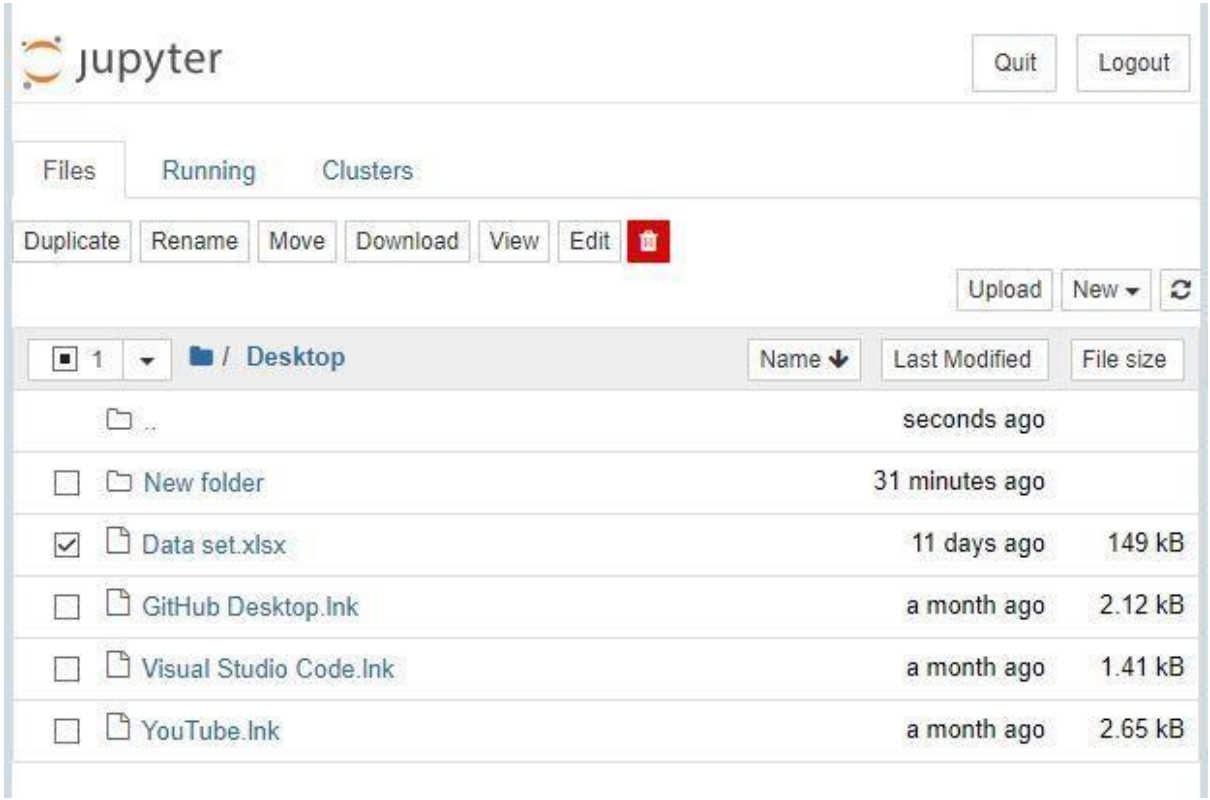
2.Load the Dataset:

- You need to have your air quality dataset in a compatible format,

such as a CSV, Excel, or other tabular format. Let's assume your dataset is in a CSV file. You can load it using the' `read_csv'` function:

```python
# Replace 'your_dataset.csv' with the actual file path of your dataset.
df = pd.read_csv('your_dataset.csv')
```

- Then click the blue Upload button displayed in the file's row to add the file to the project.

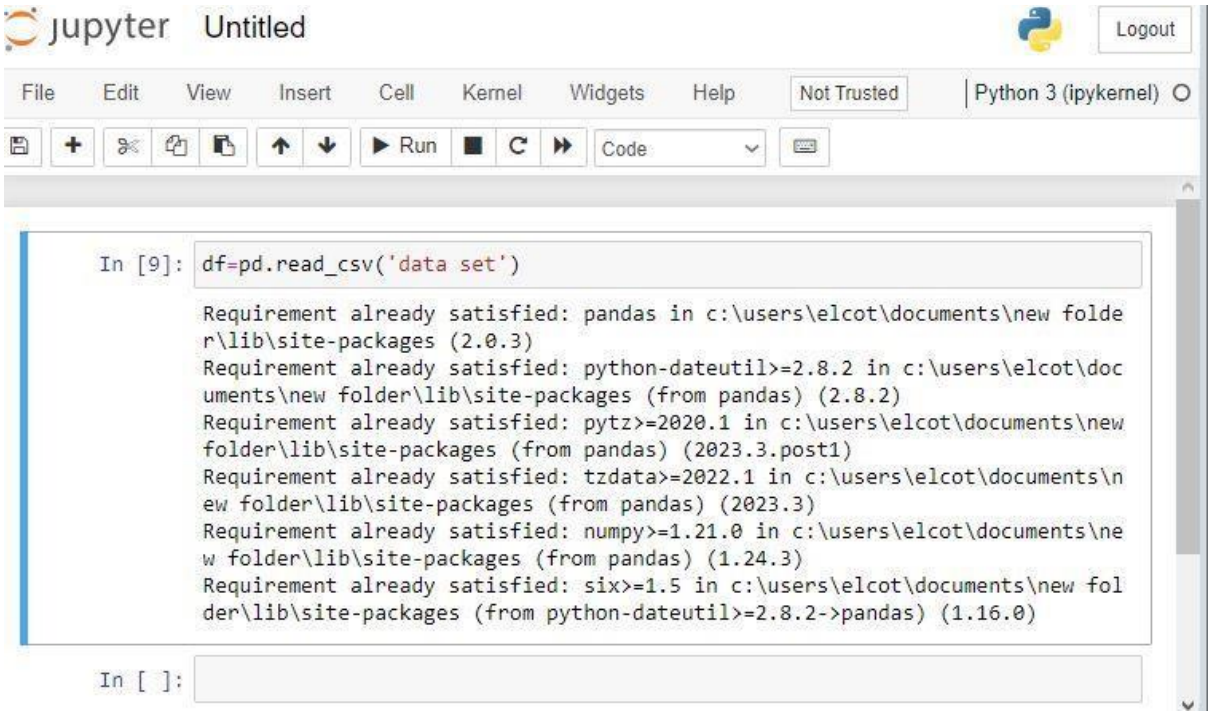


3.Explore the Dataset:

- Once you've loaded the dataset, you should explore it to understand its structure. You can start by checking the first few rows to get a glimpse of the data.

```python
# Display the first 5 rows of the dataset.
print(df.head())
```

- Explore data set the through the Object Storage Explorer and Table Explorer options on the left toolbar of the this Lab interface.



4.Data pre-processing:

- Data pre-processing is essential to ensure the data is clean and ready for analysis. Common pre-processing steps include handling missing values, removing duplicates, and dealing with data types.

Handling Missing Value:

You can check for missing values in your dataset using the ‘ isna ()’ function and then decide how to handle them. For example, to count missing values in each column.

```python
missing_values = df.isna().sum()
print(missing_values)
```

Date type conversions:

- Ensure that columns have the correct data types. For instance, dates should be converted to datetime objects if they are represented as strings.

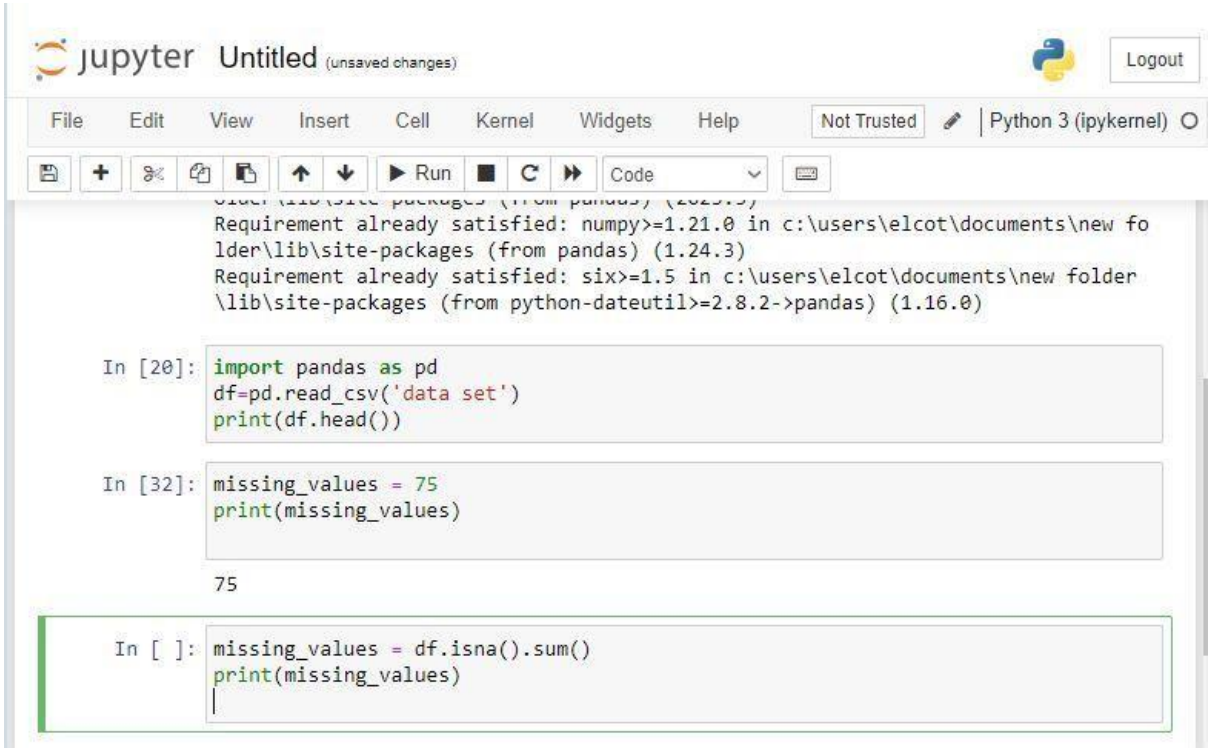
```python
# Convert a date column to datetime.
df['date_column'] = pd.to_datetime(df['date_column'])
```

Handling Duplicates:

- Check for and remove duplicates if they exist in the dataset.

python Copy code

```
df = df.drop_duplicates()
```

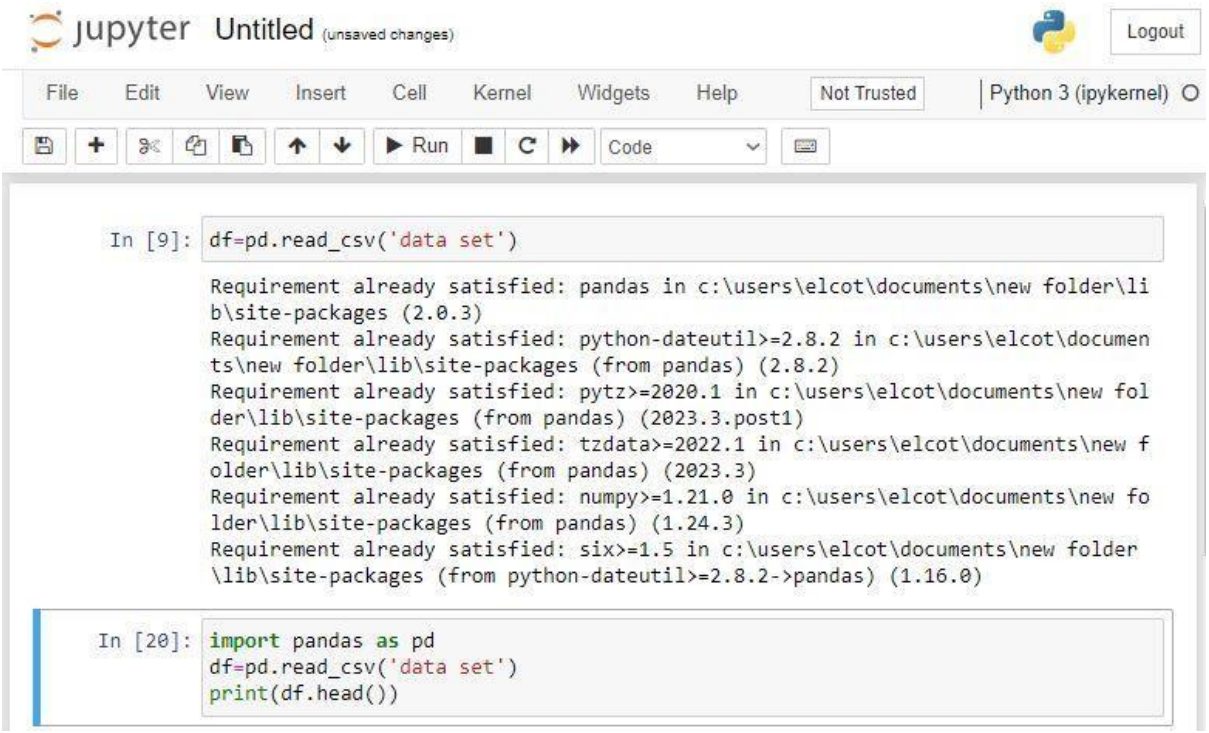


5.Save pre-processed Data:

- After pre-processing, you may want to save the pre-processed dataset for future use:

```python
df.to_csv('preprocessed_data.csv', index=False)
```

Remember to replace 'your_dataset.csv' with the actual path to your air quality dataset and adapt the pre-processing steps to your specific data. If you encounter any issues or need further assistance, feel free to ask for help with specific aspects of pre-processing.

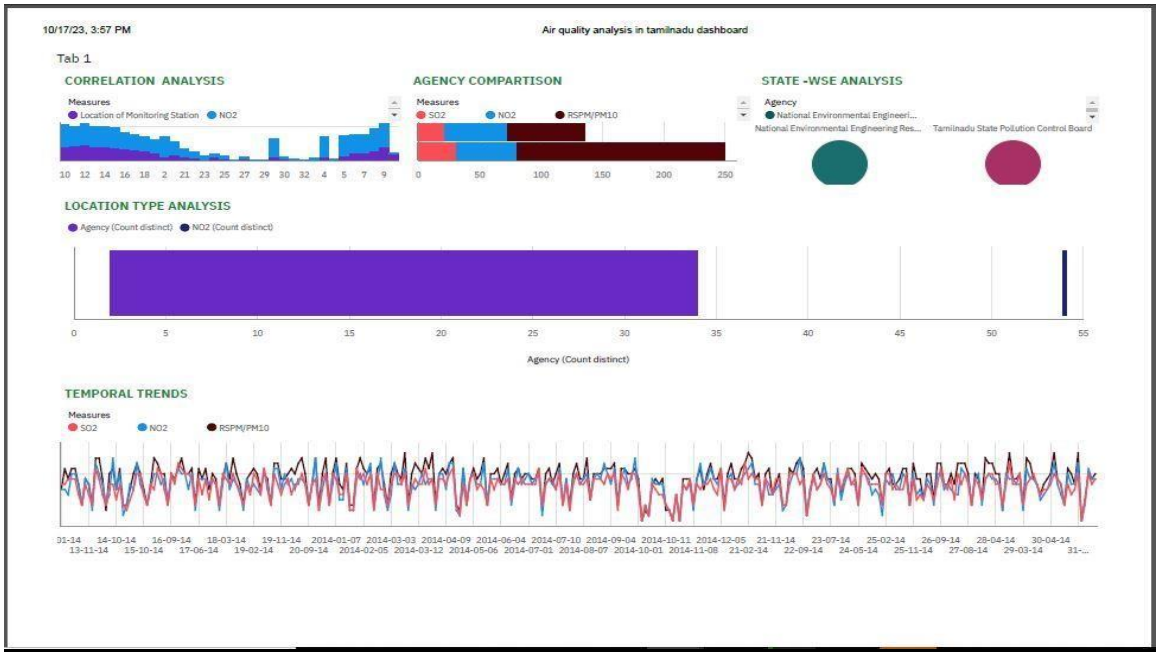


PHASE-4

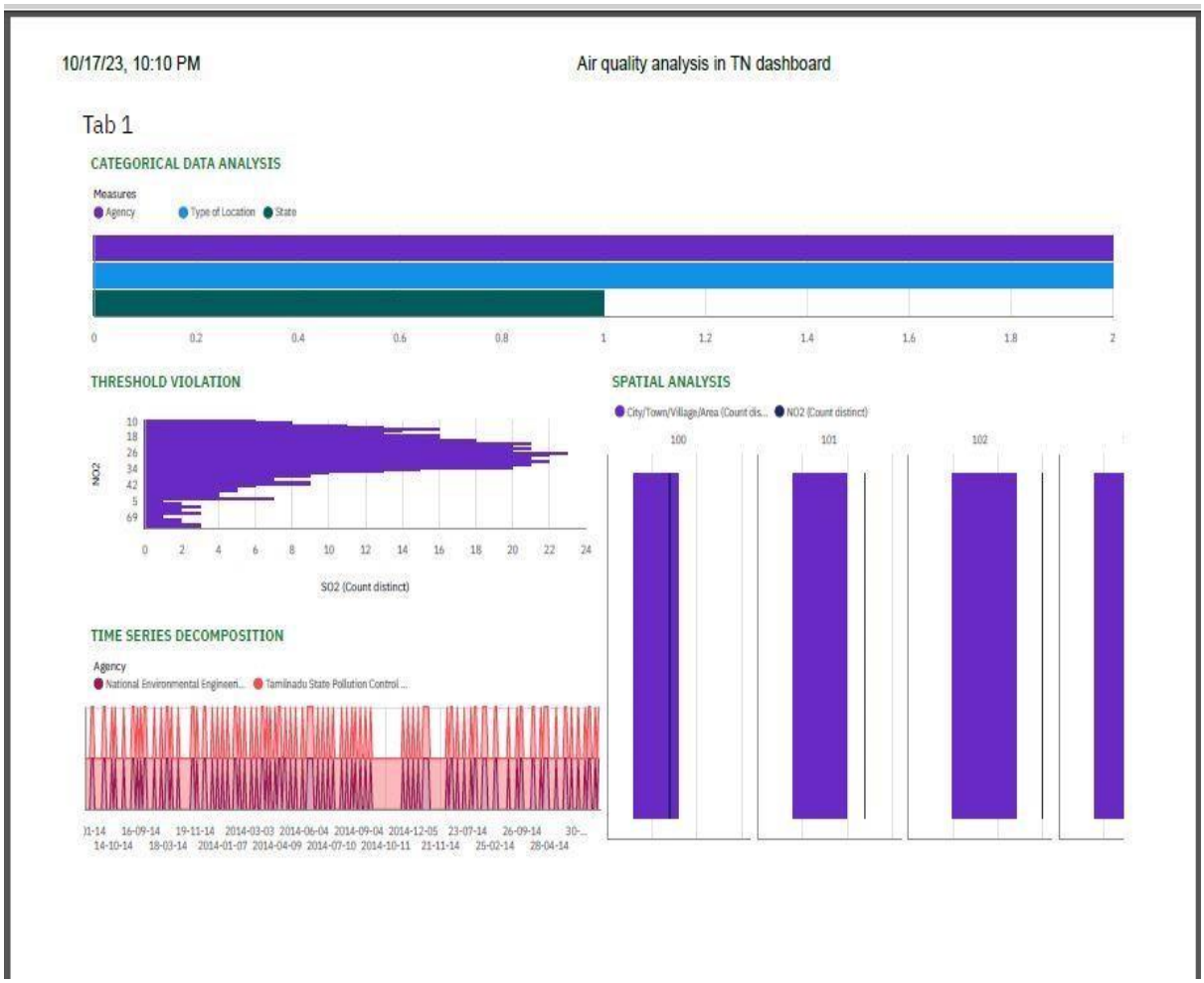
DEVELOPMENT PART-2

Create visualization using data visualization ;

- Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's storytelling with a purpose.

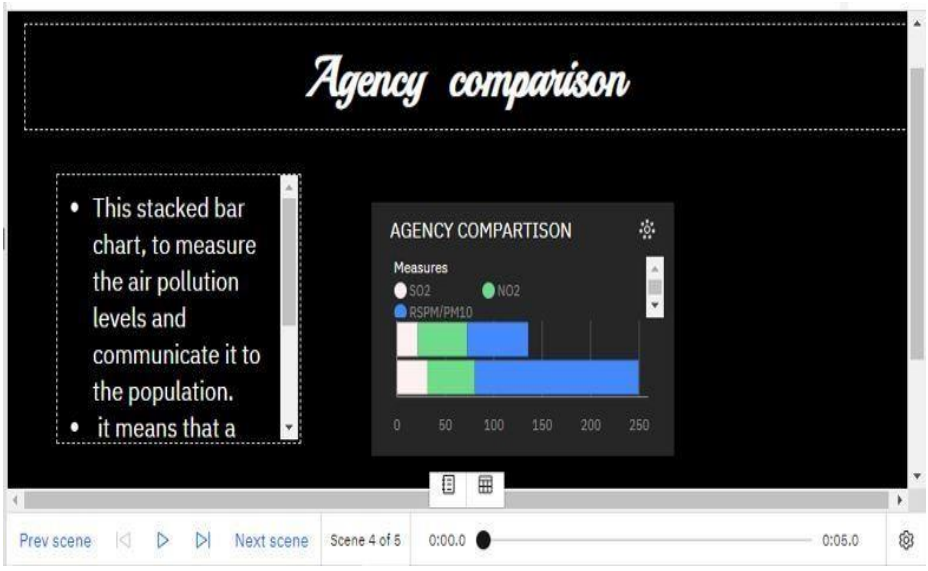


- The main goal of data visualization is **to make it easier to identify patterns, trends and outliers in large data sets**.

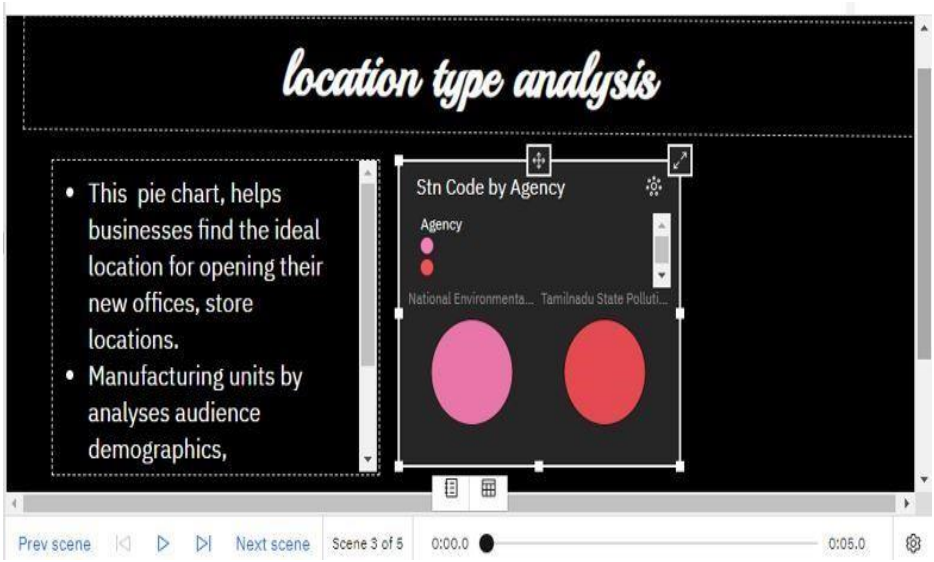


Story created by visualization:

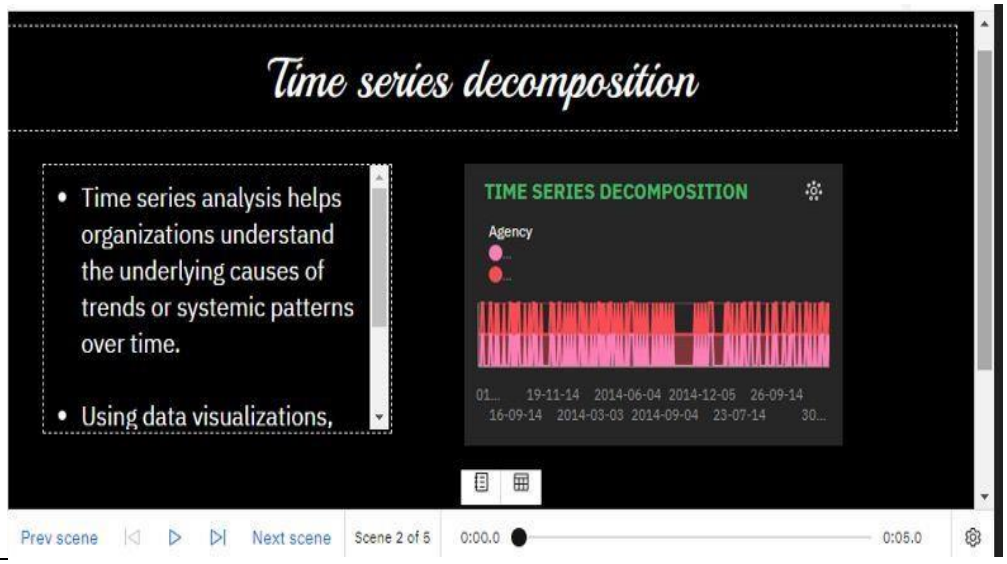
- You can assemble a view that contains visualizations, such as a graph, chart, plot, table, map, or any other visual representation of data.



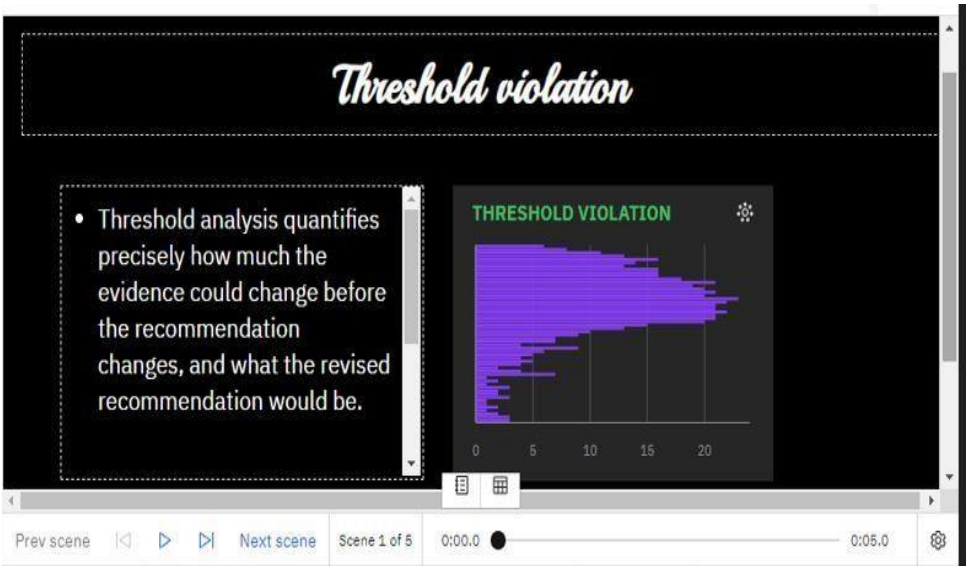
- This bar chart allows you to assess the performance of different agencies in monitoring and controlling air pollution.



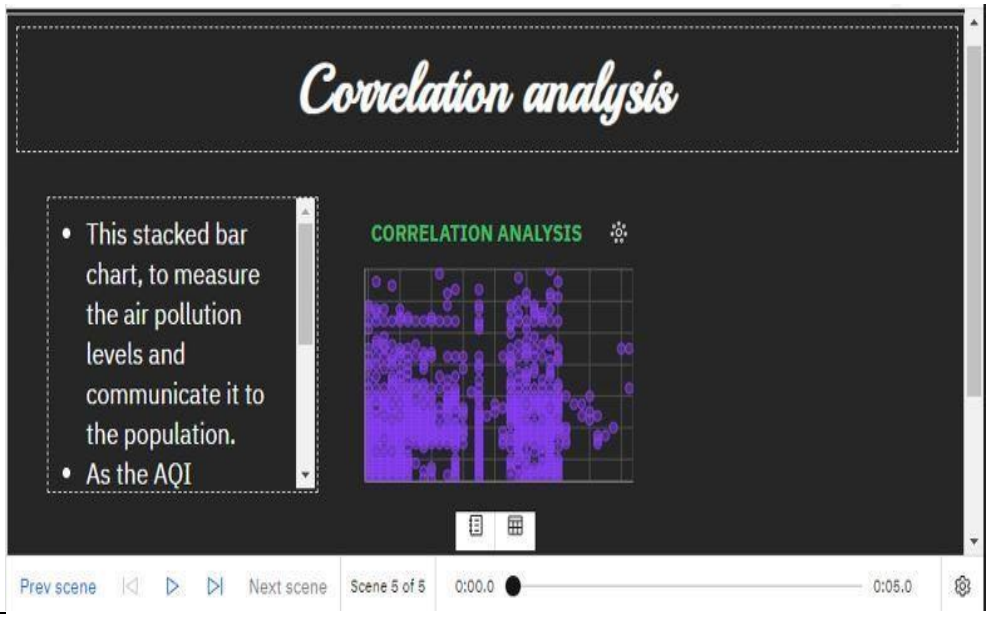
- This chart will provide a breakdown of monitoring stations by location type, helping you understand where most of the data is collected.



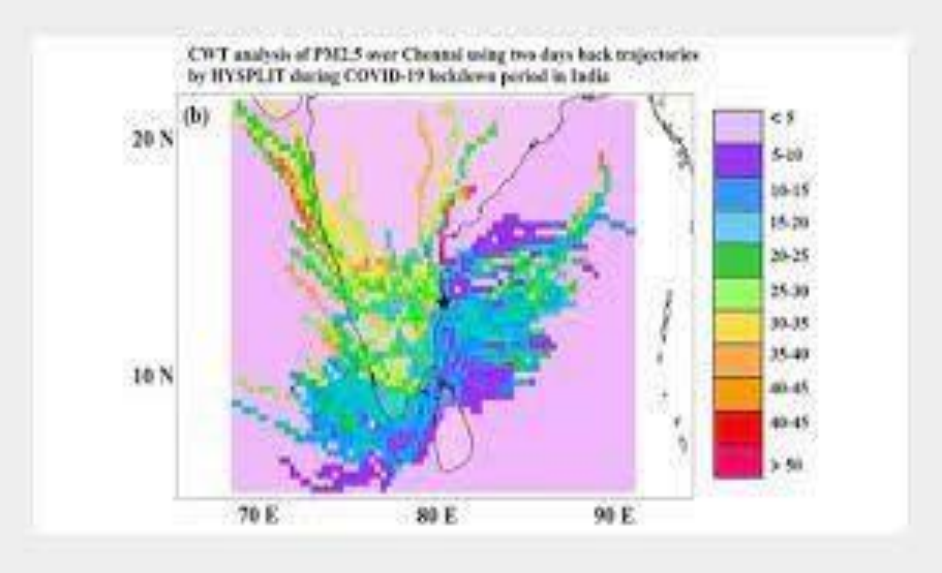
- This bar chart or map will help you compare pollutant levels across different locations, providing insights into spatial variations in air quality.

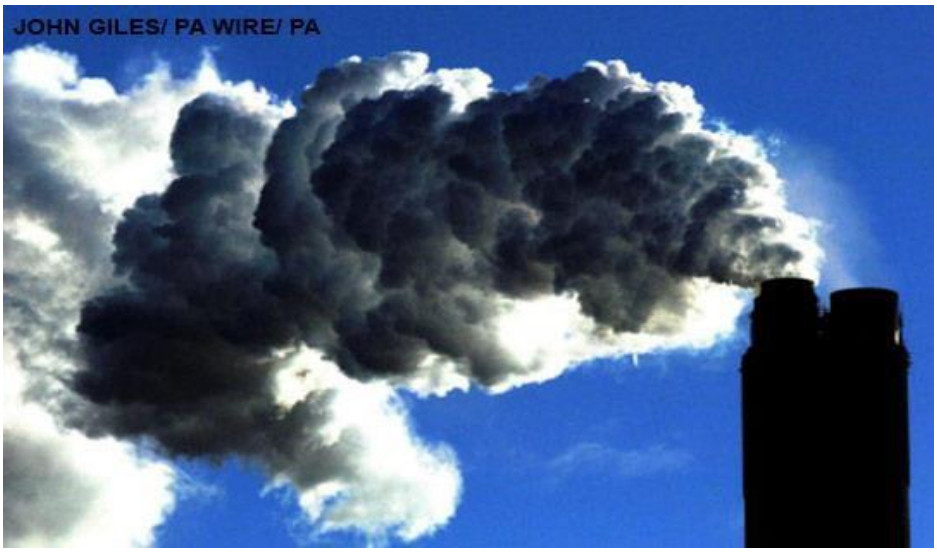


✞ This bar chart will show how frequently pollutant levels exceed regulatory thresholds or guidelines.



❖ The scatter plot will help you identify relationships and correlations between different pollutants. Are there any significant correlation.





……End of the project assessment……