AIR QUALITY INDEX ANALYZER USING ML

AN INDUSTRY ORIENTED MINI REPORT

Submitted to

**JAWAHARLAL NEHRU TECNOLOGICAL UNIVERSITY, HYDERABAD**

In partial fulfillment of the requirements for the award of the degree of

**BACHELOR OF TECHNOLOGY**

**In**

**COMPUTER SCIENCE AND ENGINEERING(AI&ML)**

Submitted by

ERRABOINA SHESHIKUMAR 21UK1A6614

BODHIREDDY NIDHI REDDY 21UK1A6658

SAKINALA SAIPRIYA 21UK1A6629

ADABOINA SHIVA 21UK1A6646

Under the guidance of

Ms. R. SWATHI

Assistant professor



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (AI AND ML)

VAAGDEVI ENGINEERING COLLEGE

BOLLIKUNTA, WARANGAL (T. S) – 506005 2021-2025

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

VAAGDEVI ENGINEERING COLLEGE

WARANGAL



CERTIFICATE OF COMPLETION

MINI PROJECT

This is to certify that the mini project report entitled by “AIR QUALITY INDEX ANALYZER USING MACHINE LEARNING” is being submitted by E.SHESHIKUMAR(21UK1A6614) , B.NIDHI(21UK1A6658) , S.SAIPRIYA(21UK1A6629) , E.SHIVA(21UK1A6646) in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in computer science and engineering(AI&ML) to Jawaharlal Nehru Technological University Hyderabad during the academic year 2024- 2025

Project guide Head of the Department

Ms. R. SWATHI Dr SHARMILA REDDY Assistant professor Professor

EXTERNAL

**ACKNOWLEDGEMENT**

We wish to take this opportunity to express our sincere gratitude and deep sense of

respect to our beloved **Dr.P.PRASAD RAO,** Principal,Vaagdevi Engineering

College for making us available all the required assistance and for his support and

inspiration to carry out this UG Project Phase-1 in the institute.

We extend our heartfelt thanks to **Dr.K.SHARMILA**, Head of the Department of

CSE, Vaagdevi Engineering College for providing us necessary infrastructure and

thereby giving us freedom to carry out the UG Project Phase-1.

We express heartfelt thanks to Smart Bridge Educational Services Private Limited,

for their constant supervision as well as for providing necessary information

regarding the UG Project Phase-1 and for their support in completing the UG Project

Phase-1.

We express heartfelt thanks to the guide,Ms. R. SWATHIAssistant professor,

Department of CSE for his constant support and giving necessary guidance for

completion of this UG Project Phase-1.

Finally, we express our sincere thanks and gratitude to my family members, friends

for their encouragement and outpouring their knowledge and experience throughout

the thesis.

ERRABOINA SHESHIKUMAR 21UK1A6614

ASAKINALA SAIPRIYA 21UK1A6629

BODHIREDDY NIDHIREDDY 21UK1A6658

EDABOINA SHIVA 21UK1A6646

**ABSTRACT**

The Air Quality Index (AQI) is a crucial tool for assessing air quality based on pollutant concentrations. It categorizes air quality into six buckets: Good, Satisfactory, Moderate, Poor, Very Poor, and Severe, with specific AQI value ranges, associated symptoms, diseases, and precautions. Good signifies excellent air quality with no notable symptoms,while Severe indicates life-threatening pollution with severe health risks.Understanding the AQI allows individuals to make informed decisions about outdoor activities and take appropriate precautions to protect their health. It serves as a valuable resource in promoting awareness and mitigating the adverse effects of air pollution on human well-being.

**TABLE OF CONTENTS:-**

**1.INTRODUCTION .................................…..................…....………… 5**

**1.1 OVERVIEW.............................................….............…………….... 5**

**1.2 PURPOSE .................................…...............……............................. 5**

**LITERATURE SURVEY ...........................………........…................... 8**

**2.1 EXISTING PROBLEM ....................……....................................... 8**

**2.2 PROPOSED SOLUTION..............................................……….....8-9**

**THEORITICAL ANALYSIS… ..................................………............ 10**

**3.1 BLOCK DIAGRAM..................................................…….............. 10**

**3.2 HARDWARE /SOFTWARE DESIGNING .............………....10-11**

**EXPERIMENTAL INVESTIGATIONS....................………....... 12-13**

**FLOWCHART…..............................................................………........ 14**

**RESULTS… .............................................................……….….......15-18**

**ADVANTAGES AND DISADVANTAGES…..........………............. 19**

**APPLICATIONS ....................................................………................. 20**

**CONCLUSION ..................................................………...................... 20**

**FUTURE SCOPE… .........................................………….................... 21**

**BIBILOGRAPHY..............................................…………...............22-23**

1. **INTRODUCTION**
   1. OVERVIEW

Clean and healthy air is a fundamental necessity for human well-being, and the Air Quality Index (AQI) serves as a critical tool for assessing and understanding the quality of the air we breathe. The AQI system categorizes air quality into six distinct buckets, each defined by specific AQI value ranges. These categories, ranging from "Good" to "Severe," provide essential information about the levels of pollutants in the atmosphere and their potential impacts on our health. This introduction explores the AQI classification, the associated symptoms and diseases, as well as the recommended precautions at each level, emphasizing the importance of AQI awareness in safeguarding public health. Understanding the AQI is vital in making informed decisions about outdoor activities and taking proactive measures to minimize the risks associated with varying air quality conditions. In this overview, we delve into the AQI classification, the associated symptoms and diseases at each level, and the recommended precautions to protect public health. Recognizing the significance of AQI awareness is essential in making informed decisions about outdoor activities and implementing preventive measures to mitigate health risks associated with fluctuating air quality conditions. It is a fundamental aspect of safeguarding the well-being of individuals and communities in an increasingly polluted world.

* 1. **PURPOSE**

The Air Quality Index (AQI) serves several critical purposes in assessing and communicating air quality, with a primary focus on safeguarding public health and the environment. In detail, its purposes include:

1. **Informing the Public**: The AQI is designed to provide the general public with easily understandable information about air quality. It informs individuals about the safety of outdoor activities and helps them make informed decisions to protect their health.

2. **Health Protection**: One of its primary purposes is to protect public health. The AQI categorizes air quality into different levels, each associated with specific health risks. This enables individuals, particularly those with respiratory conditions, to take precautions based on the current air quality conditions.

**3. Government Regulation**: The AQI is used by regulatory agencies and governments to set air quality standards and policies. It helps in identifying areas with poor air quality and implementing measures to improve it, such as emission controls and pollution reduction strategies.

1. **Environmental Monitoring**: The AQI also plays a role in monitoring the impact of air pollution on the environment. It helps track changes in air quality over time and assess the effectiveness of pollution control measures.
2. **Economic Impact**: Understanding air quality is essential for evaluating the economic impact of pollution on industries, healthcare costs, and productivity. It can guide policy decisions that affect economic development.
3. **LITERATURE SURVEY**

**2.1 EXISTING PROBLEM**

➢ The existing problem of air quality degradation is a multifaceted global challenge that spans health, environmental, and economic domains. Poor air quality, characterized by elevated levels of pollutants, takes a devastating toll on public health, contributing to a wide array of respiratory and cardiovascular diseases while being a leading cause of premature mortality.

➢ Moreover, air pollution wreaks havoc on ecosystems, causing harm to both flora and fauna, and contributes to environmental issues like acid rain and smog. Economically, it translates into substantial costs related to healthcare expenses, lost workdays, and decreased labor productivity. Additionally, air pollution plays a significant role in climate change, with greenhouse gas emissions intensifying global warming. Vulnerable communities, often located near pollution sources, bear a

disproportionate burden.

➢ Inadequate regulations, limited public awareness, and the variability of air quality further compound the issue. Addressing these problems necessitates stringent regulations, the transition to cleaner energy sources, robust public education, and international cooperation to combat the health, environmental, and economic consequences associated with poor air quality.

**2.2 PROPOSED SOLLUTION**

➢ Our innovative proposed solution leverages advanced predictive modeling to anticipate and communicate Air Quality Index (AQI) bucket categories accurately and efficiently. By forecasting AQI buckets, we enable proactive dissemination of information regarding air quality, accompanying health risks, and recommended precautions.

This solution integrates technology, data analytic s, and public health awareness to create a comprehensive system for managing air quality issues.

1. **Advanced Predictive Modeling**: We employ sophisticated predictive models that take into account various environmental factors, historical data, and real time monitoring to forecast AQI buckets. This ensures timely and reliable information for the public.

2. **Customized Health Information**: For each AQI category, our system

provides a tailored list of potential diseases and recommended precautions. This empowers individuals to take proactive steps to protect their health based on the fore castes air quality conditions.

1. **User-Friendly Interfaces**: Our solution offers user-friendly interfaces, making it easy for the public to access and understand AQI predictions and associated health information.
2. **Community Engagement**: We actively engage with communities to raise awareness about the AQI forecasting system, ensuring that the public is well informed and prepared to take necessary precautions.

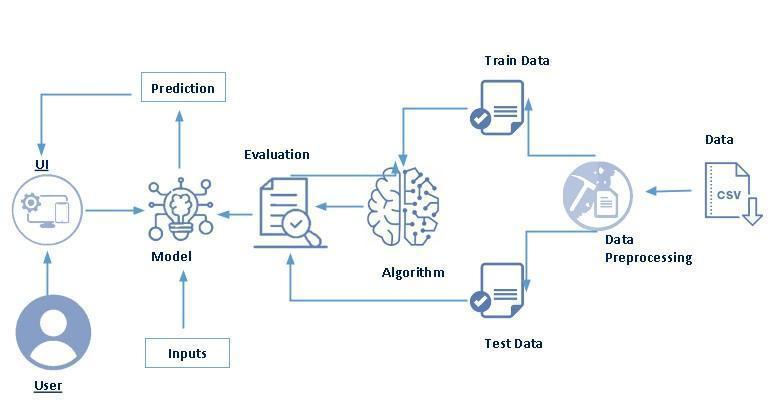
5. **Educational Initiatives**: Our solution includes educational initiatives to raise awareness about the importance of air quality and the impact of air pollution on public health, empowering individuals to make informed choices

➢ By accurately predicting AQI bucket categories and providing associated health information, our solution empowers individuals and communities to make informed decisions, protect their health, and reduce the adverse health effects of air pollution. It is a comprehensive approach to proactively address air quality issues and create healthier environments for all.

➢ By encompassing these elements, our proposed solution creates a comprehensive and adaptable framework for AQI prediction and management, with a strong emphasis on accuracy, user-friendliness, and community engagement.

**3.THEORITICAL ANALYSIS**

**3.1. BLOCK DIAGRAM**



**3.2. SOFTWARE DESIGNING**

The following is the Software required to complete this project:

➢ **Jupyter Notebook:** Jupyter Notebook will serve as the development and execution environment for your predictive modeling, data preprocessing, and model training tasks. It provides an interactive, web-based environment with access to Python libraries and hardware acceleration.

➢ **Dataset (CSV File):** The dataset in CSV format is essential for training and testing your predictive model. It should include historical air quality data, weather information, pollutant levels, and other relevant features.

➢ **Data Preprocessing Tools**: Python libraries like NumPy, Pandas, and Scikit-learn will be used to preprocess the dataset. This includes handling missing data, feature scaling, and data cleaning.

➢ **Feature Selection/Drop**: Feature selection or dropping unnecessary features from the dataset can be done using Scikit-learn or custom Python code to enhance the model's efficiency.

➢ **Model Training Tools**: Machine learning libraries such as Scikit-learn,

TensorFlow, or PyTorch will be used to develop, train, and fine-tune the predictive model. Regression or classification models can be considered, depending on the nature of the AQI prediction task.

➢ **Model Performance Evaluation:** After model training, performance evaluation tools, such as Scikit-learn metrics or custom validation scripts, will assess the model's predictive capabilities. You'll measure the model's ability to predict AQI categories based on historical data using the R² score.

➢ **UI Based on Flask Environment**: Flask, a Python web framework, will be used to develop the user interface (UI) for the system. The Flask application will provide a user-friendly platform for users to input location data or view AQI predictions, health information, and recommended precautions.

➢ Jupyter Notebook will be the central hub for model development and training, while Flask will facilitate user interaction and data presentation. The dataset, along with data preprocessing, will ensure the quality of the training data, and feature selection will optimize the model. Finally, model performance evaluation will confirm the system's predictive capabilities, allowing users to rely on the AQI predictions and associated health information.

**4.EXPERIMENTAL INVESTIGATION**

In this project, we have used Air Quality Dataset. This dataset is a csv file consisting of labelled data and having the following columns-

1. **City**: Location or urban area for which air quality data is recorded.

2. **Date**: The specific date on which air quality measurements were taken.

3. **PM2.5**: Particulate Matter with a diameter of 2.5 micrometers, a key air pollutant.

4. **PM10**: Particulate Matter with a diameter of 10 micrometers, another significant air pollutant.

5. **NO**: Nitric Oxide, a gaseous air pollutant.

6. **NO2**: Nitrogen Dioxide, a toxic gas often related to combustion processes.

7. **NOx**: Nitrogen Oxides, a group of nitrogen-containing air pollutants.

8. **NH3**: Ammonia, a compound that can contribute to air pollution.

9. **CO**: Carbon Monoxide, a colorless, odorless gas produced by incomplete combustion.

10.**SO2**: Sulfur Dioxide, a toxic gas often linked to industrial processes.

11.**O3**: Ozone, a secondary pollutant with varying health impacts.

12.**Benzene**: A volatile organic compound that can be harmful when inhaled.

13.**Toluene**: Another volatile organic compound commonly found in urban air.

14.**Xylene**: A group of volatile organic compounds, often associated with vehicle emissions.

15.**AQI**: Air Quality Index, a numerical value indicating overall air quality.

16.**AQI\_Bucket**: The qualitative classification of air quality based on the AQI value.

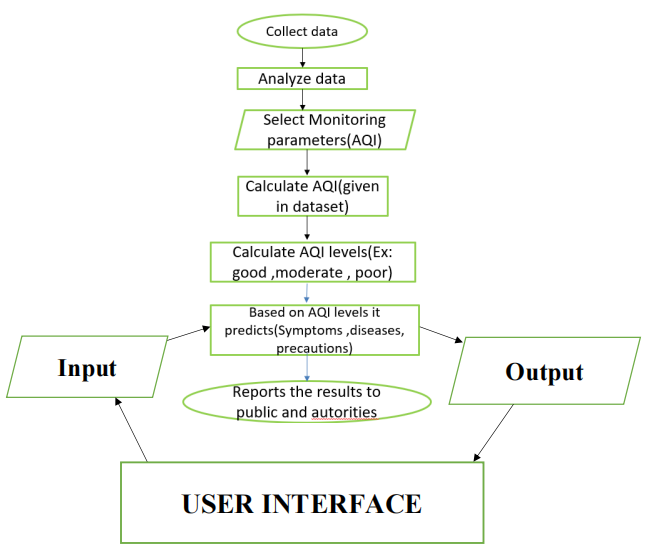
For the dataset we selected, it consists of more than the columns we want to predict it .

So, we have chosen the feature drop it contains the columns that we are going to predict

the AQI value.

➢ Feature drop means it drops the columns that we don’t want in our dataset.

1. **FLOWCHART**



1. **RESULT**

**ABOUT PAGE**



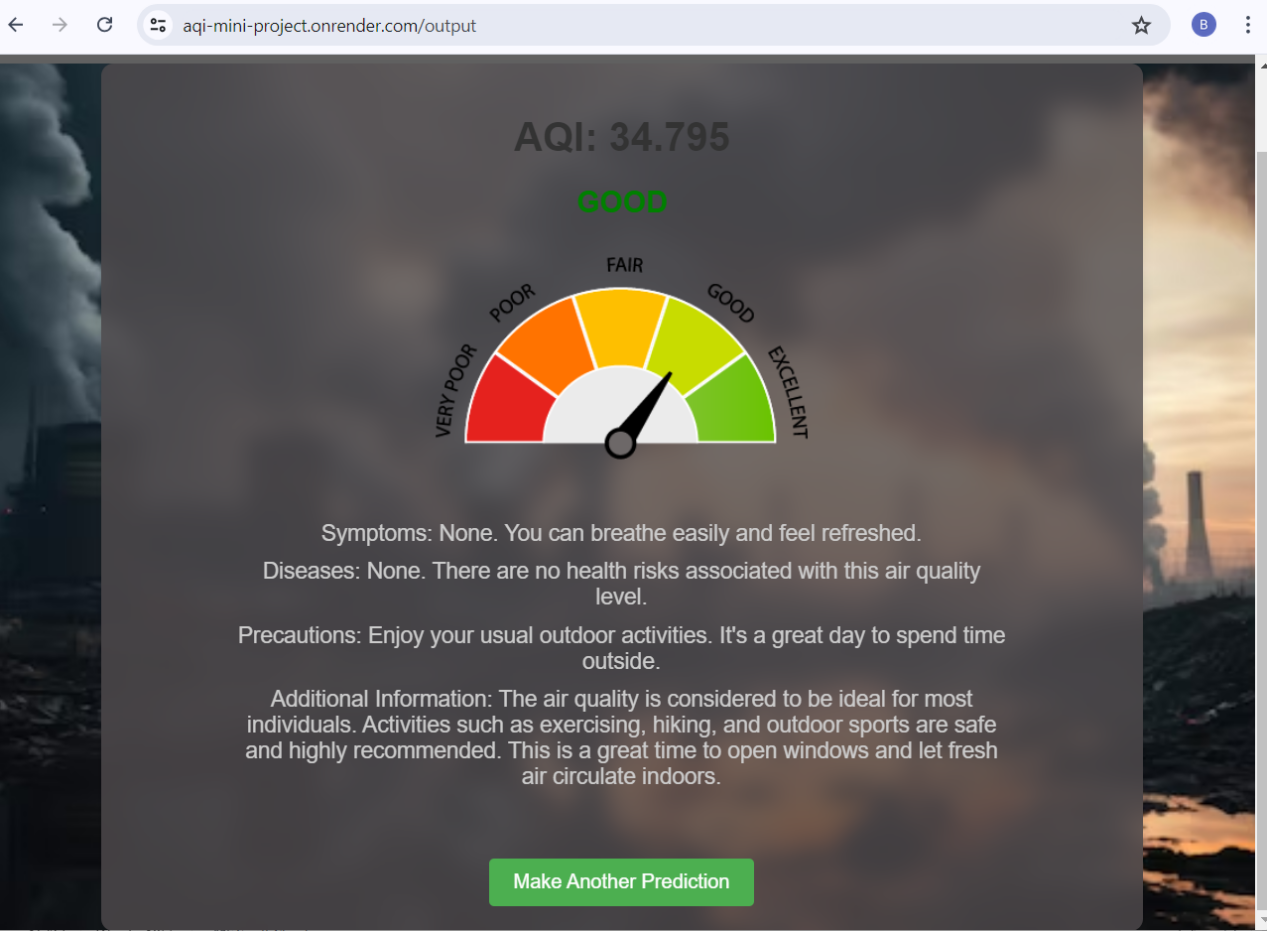
**HOME PAGE**



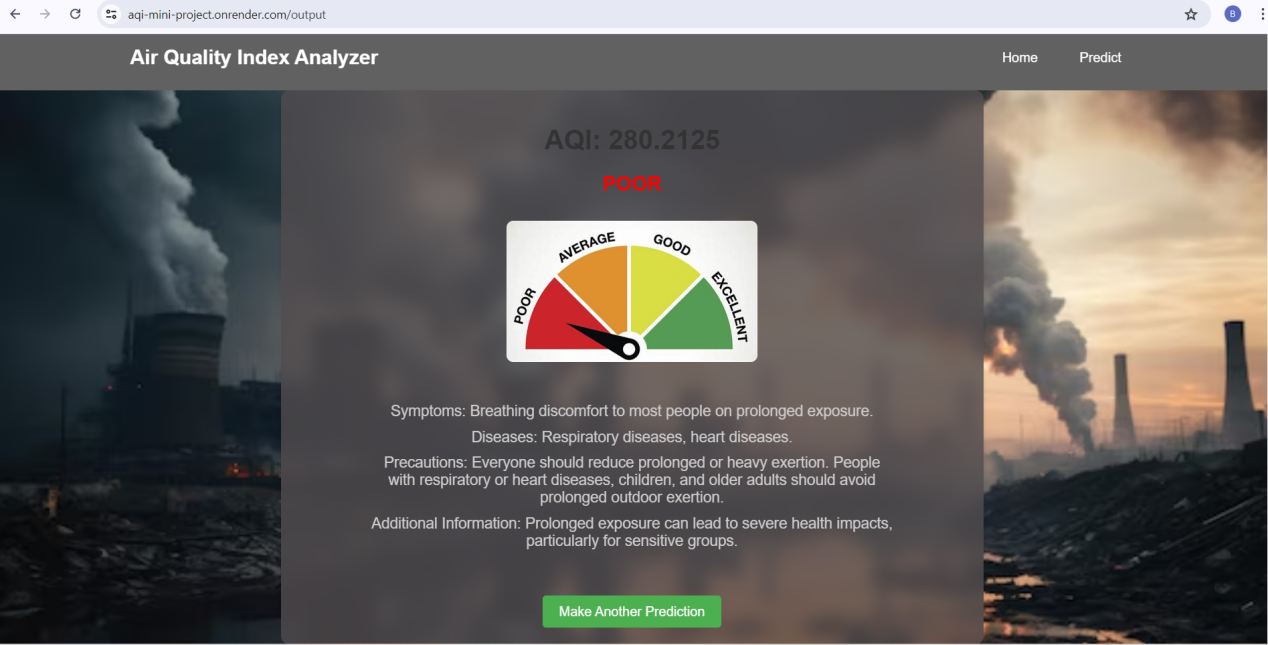
**PREDICTIONS**



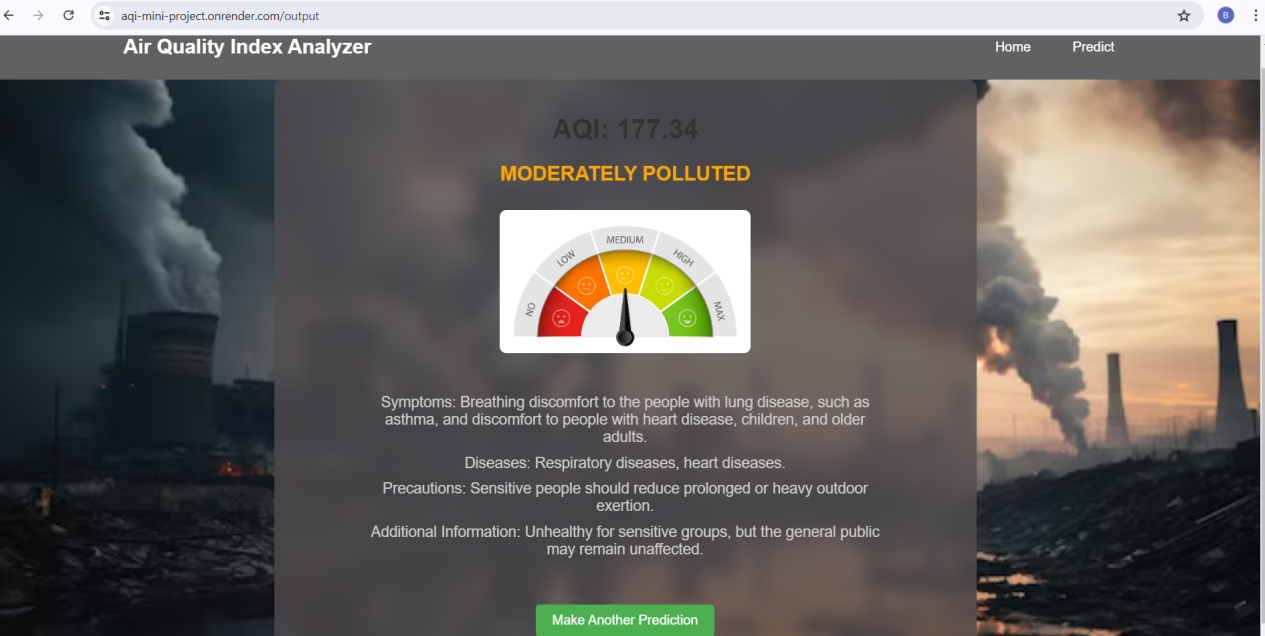
**OUTPUT 1**

****

**OUTPUT 2**

****

**OUTPUT 3**

****

**7.ADVANTAGES AND DISADVANTAGES**

**ADVANTAGES:**

1. **Enhanced Public Health**: Provides timely air quality information, promoting health protection.

2. **Proactive Decision-Making**: Enables informed decisions regarding outdoor activities.

3. **Environmental Awareness**: Increases awareness of air pollution and its impacts.

4. **Community Engagement**: Engages communities in air quality management and feedback.

5. **Customized Health Information**: Tailors health recommendations to AQI categories.

**DISADVANTAGES:**

1. **Data Reliance**: Relies on accurate and up-to-date air quality data for precise predictions.

2. **Resource-Intensive**: Requires substantial resources for data collection, modeling, and system maintenance.

3. **Technical Expertise**: Users may need some technical knowledge to interpret AQI and health information.

4. **Model Accuracy**: The system's accuracy depends on the quality of the predictive model.

5. **Privacy Concerns**: User data may raise privacy concerns and necessitate secure data handling.

**8.APPLICATIONS**

1. **Public Health Protection**: Empowering individuals to make informed decisions regarding outdoor activities, reducing exposure to poor air quality, and minimizing health risks.

2. **Environmental Monitoring**: Assessing the impact of air pollution on the environment, ecosystems, and natural habitats, aiding in conservation efforts.

3. **Government Policy**: Assisting governments and regulatory bodies in setting air quality standards, formulating pollution control policies, and conducting effective urban planning

4. **Public Awareness**: Raising public awareness about the importance of air quality and its impact on health, influencing behavior and lifestyle choice.

**9.CONCLUSION**

➢ In conclusion, the proposed Air Quality Index (AQI) prediction and management system presents a holistic solution to address the critical issues of air quality monitoring, public health protection, and community engagement. By integrating advanced predictive modeling with user-friendly interfaces, the system empowers individuals to make informed decisions, safeguard their health, and actively participate in air quality management. The project's key components, including data preprocessing, feature selection, model training, and user interface development, create a comprehensive framework for delivering real-time AQI predictions and associated health information.

➢ In the ever-evolving field of environmental science and technology, this project represents a significant step forward in ensuring cleaner air and healthier living environments. With ongoing research, innovation, and community involvement, the system has the potential to make a positive impact on public health, environmental conservation, and global collaboration in mitigating air pollution issues

**10.FUTURE SCOPE**

Future Scope of the AQI Prediction and Management System:

1. **Global Expansion**: Extend the system's reach to more regions and countries, addressing air quality issues on a global scale.

2. **Advanced Technology Integration**: Integrate IoT sensor networks and smart city initiatives for real-time air quality monitoring and urban planning.

3. **Air Quality Forecasting**: Enhance the system's capabilities for short and long term air quality forecasting.

4. **Healthcare Integration**: Collaborate with healthcare providers to incorporate AQI information into patient care, particularly for those with respiratory conditions, improving public health outcomes.

**11.BIBILOGRAPHY**

[1] Anderson H. R., R.W. Atkinson, J. L. Peacock, M. J. Sweeting and L. Marston. Ambient Particulate matter and health effect;Publication bias in studies of short-term association. Epidemiol 16; 2005: 155- 163.

[2] Analitis A.,K. Katsouyanni, E. Dimakopoulou, A. K. Samoli,Y.Nikolouopoulos, G. Petasakis, J. Touloumi, H. Schwartz, H. R.Anderson, K. Cambra, F. Forastiere, D. Zmirou, J. M. Vonk,L.clancy, B. Kriz, J. Bobvos and J. Pekkanen. Short-term effects of ambient paricles on cardiovascular and resipiratory mortilaity. Epidemiol 17; 2006: 230-233.

[3] Kumar A.,Goyal P. Forecasting of air quality in Delhi using principal component regression technique. Atmospheric Pollution Research. 2 . 2011: 436-444.

[4] Central Pollution Control Board (CPCB).Guidelines for National ambient air quality monitoring, Series:NAAQM/25/2003-04.Parivesh Bhavan,Delhi; 2009:

[5] Central pollution control board (CPCB). National air quality index , Series CUPS/82/2014-15; 2014:

[6] Ekpenyong E. C.,Eltebong E. O., Akpan E. E., Samson T. K.,Danierl E. N. Urban city transportation mode and respiratory health effect of an air pollution: a cross sectional study among transit and non transit worker in Nigeria. BMJ open,doi:10.1136/bmjopen-2012-001253; 2012:

[7] Kaushik. C. P., Ravindra K., Yadav K., Mehta S. and Haritash A.K.. Assessment of ambient air quality in urban centres of haryana(India) in relation to different anthropogenic activities and health risks. Environment Monitoring and Assement. 122; 2006:27-40.

[8] Pipalatkar. P. P. , Gajghate. D.G and Khaparde V.V. Source identification of different size fraction of PM10 Using Factor analysis at residential cum commercial Area of Nagpur city. Bull.Envionment Contam Toxicol. 88;2012: 260-264.

[9] Pope C. A. III and D. W.Dockery Health effects of fine particulate air pollution; lines that connect. Journal of Air and Waste Management Association. 56; 2006: 709-742.

[10] Bhuyan P. K.,Samantray P., Rout S. P. Ambient air quality status in choudwar area of cuttack district.International Journal of Environmental Sciences. 1; 2010:

[11] Ravikumar,P., Prakash, K.L. and Somashekar,R.K.. Air quality Indices to understand the ambient air quality in the vicity of dam site of different irrigation projects in karanataka state, india.International 22journal of science and nature. 5; 2014 : 531-541.

[12] U. S. Environmental Protection Agency (USEPA). Guidelines for reporting of daily air quality- air quality index (AQI), Series EPA-454/B-06-001. Research Trangle Park , North carolina.2006

**12.APPENDIX**

**Model building :**

1)Dataset

2)Jupyter Notebook and Spyder Application Building

1. HTML file (Index file,Home file, Predict file,Output file )

1. CSS file

2. Models in pickle format

**SOURCE CODE:**

**INDEX.HTML**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="utf-8" />

<meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no" />

<meta name="description" content="" />

<meta name="author" content="" />

<title>AQI Analyzer</title>

<link rel="icon" type="image/x-icon" href="assets/favicon.ico" />

<!-- Font Awesome icons (free version)-->

<script src="https://use.fontawesome.com/releases/v6.3.0/js/all.js" crossorigin="anonymous"></script>

<!-- Google fonts-->

<link href="https://fonts.googleapis.com/css?family=Varela+Round" rel="stylesheet" />

<link href="https://fonts.googleapis.com/css?family=Nunito:200,200i,300,300i,400,400i,600,600i,700,700i,800,800i,900,900i" rel="stylesheet" />

<!-- Core theme CSS (includes Bootstrap)-->

<link href="{{ url\_for('static', filename='styles.css') }}" rel="stylesheet" />

<style>

.card {

margin: 10px; /\* Adjust the value as needed \*/

}

.card-bottom-space {

margin-bottom: 20px; /\* Adjust the value as needed \*/

padding-bottom: 20px; /\* Adjust the value as needed \*/

}

</style>

</head>

<body id="page-top">

<!-- Navigation-->

<nav class="navbar navbar-expand-lg navbar-light fixed-top" id="mainNav">

<div class="container px-4 px-lg-5">

<a class="navbar-brand" href="#page-top">AQI Analyzer</a>

<button class="navbar-toggler navbar-toggler-right" type="button" data-bs-toggle="collapse" data-bs-target="#navbarResponsive" aria-controls="navbarResponsive" aria-expanded="false" aria-label="Toggle navigation">

Menu

<i class="fas fa-bars"></i>

</button>

<div class="collapse navbar-collapse" id="navbarResponsive">

<ul class="navbar-nav ms-auto">

<li class="nav-item"><a class="nav-link" href="{{url\_for('predict')}}">Predict</a></li>

<li class="nav-item"><a class="nav-link" href="#about">About</a></li>

<li class="nav-item"><a class="nav-link" href="#team">Team</a></li>

</ul>

</div>

</div>

</nav>

<header class="masthead">

<div class="container px-4 px-lg-5 d-flex h-100 align-items-center justify-content-center">

<div class="d-flex justify-content-center">

<div class="text-center">

<h1 class="mx-auto my-0 text-uppercase">Air Quality Index Analyzer</h1>

<h2 class="text-white-50 mx-auto mt-2 mb-5">Your go-to platform for air quality analysis and predictions.</h2>

<a class="btn btn-primary" href="{{ url\_for('home') }}">Get Started</a>

</div>

</div>

</div>

</header>

<!-- About-->

<section class="about-section text-center" id="about">

<div class="container px-4 px-lg-5">

<div class="row gx-4 gx-lg-5 justify-content-center">

<div class="col-lg-8">

<h2 class="text-white mb-4">Real Time Air Quality India</h2>

<p class="text-white-50">

The Air Quality Index (AQI) is a measure used to communicate how polluted the air currently is or how polluted it is forecast to become. The AQI is calculated based on the levels of various pollutants in the air, including particulate matter (PM10 and PM2.5), ground-level ozone (O3), carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2)

<a href="https://www.aqi.in/dashboard/india">Real time Overview.</a>

Monitoring and understanding the AQI is crucial for public health as it provides essential information on the safety and quality of the air we breathe. High levels of air pollution can cause a variety of health problems, particularly for vulnerable groups such as children, the elderly, and those with respiratory or cardiovascular conditions. The AQI helps individuals make informed decisions about outdoor activities and take necessary precautions to protect their health.

</p>

</div>

</div>

</div>

</section>

<!-- Team-->

<section class="about-section text-center" id="team">

<div class="container px-4 px-lg-5">

<div class="row gx-4 gx-lg-5 justify-content-center">

<div class="col-lg-8">

<h2 class="text-white mb-4">Our Team</h2>

<div class="row gx-4 gx-lg-5">

<div class="col-md-6 mb-3 mb-md-0">

<div class="card py-4 h-100">

<div class="card-body text-center">

<h4 class="text-uppercase m-0">Sheshikumar</h4>

<hr class="my-4 mx-auto" />

<div class="small text-black-50">21UK1A6614</div>

</div>

</div>

</div>

<div class="col-md-6 mb-3 mb-md-0">

<div class="card py-4 h-100">

<div class="card-body text-center">

<h4 class="text-uppercase m-0">Sai Priya</h4>

<hr class="my-4 mx-auto" />

<div class="small text-black-50">21UK1A29</div>

</div>

</div>

</div>

<div class="col-md-6 mb-3 mb-md-0">

<div class="card py-4 h-100">

<div class="card-body text-center">

<h4 class="text-uppercase m-0">Nidhi Reddy</h4>

<hr class="my-4 mx-auto" />

<div class="small text-black-50">21UK1A58</div>

</div>

</div>

</div>

<div class="col-md-6 mb-3 mb-md-0">

<div class="card py-4 h-100">

<div class="card-body text-center">

<h4 class="text-uppercase m-0">Shiva</h4>

<hr class="my-4 mx-auto" />

<div class="small text-black-50">21UK1A6646</div>

</div>

</div>

</div>

</div>

</div>

</div>

</div>

</section>

<!-- Footer-->

<footer class="footer bg-black small text-center text-white-50">

<div class="container px-4 px-lg-5">Copyright &copy; Your Website 2023</div>

</footer>

<!-- Bootstrap core JS-->

<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.3/dist/js/bootstrap.bundle.min.js"></script>

<!-- Core theme JS-->

<script src="js/scripts.js"></script>

<!-- SB Forms JS-->

<script src="https://cdn.startbootstrap.com/sb-forms-latest.js"></script>

</body>

</html>

**PREDICT.HTML**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<title>Predict AQI</title>

<link rel="stylesheet" href="{{ url\_for('static', filename='predict.css') }}">

</head>

<body>

<header>

<div class="header-container">

<h1>Air Quality Index Analyzer</h1>

<nav>

<ul>

<li><a href="{{ url\_for('index') }}">Home</a></li>

<li><a href="{{ url\_for('predict') }}">Predict</a></li>

</ul>

</nav>

</div>

</header>

<div class="container">

<h1>Enter Air Quality Data</h1>

<form action="{{ url\_for('output') }}" method="POST">

<select id="city" name="city" required>

<option value="" disabled selected>Select City</option>

<option value="Ahmedabad">Ahmedabad</option>

<option value="Aizawl">Aizawl</option>

<option value="Amaravati">Amaravati</option>

<option value="Amritsar">Amritsar</option>

<option value="Bengaluru">Bengaluru</option>

<option value="Bhopal">Bhopal</option>

<option value="Brajrajnagar">Brajrajnagar</option>

<option value="Chandigarh">Chandigarh</option>

<option value="Chennai">Chennai</option>

<option value="Coimbatore">Coimbatore</option>

<option value="Delhi">Delhi</option>

<option value="Ernakulam">Ernakulam</option>

<option value="Gurugram">Gurugram</option>

<option value="Guwahati">Guwahati</option>

<option value="Hyderabad">Hyderabad</option>

<option value="Jaipur">Jaipur</option>

<option value="Jorapokhar">Jorapokhar</option>

<option value="Kochi">Kochi</option>

<option value="Kolkata">Kolkata</option>

<option value="Lucknow">Lucknow</option>

<option value="Mumbai">Mumbai</option>

<option value="Patna">Patna</option>

<option value="Shillong">Shillong</option>

<option value="Talcher">Talcher</option>

<option value="Thiruvananthapuram">Thiruvananthapuram</option>

<option value="Visakhapatnam">Visakhapatnam</option>

</select>

<input type="number" id="pm25" name="pm25" placeholder="PM2.5" step="0.001" required>

<input type="number" id="pm10" name="pm10" placeholder="PM10" step="0.001" required>

<input type="number" id="no2" name="no2" placeholder="NO2" step="0.001" required>

<input type="number" id="co" name="co" placeholder="CO" step="0.001" required>

<input type="number" id="so2" name="so2" placeholder="SO2" step="0.001" required>

<input type="number" id="o3" name="o3" placeholder="O3" step="0.001" required>

<input type="date" id="date" name="date" required>

<button type="submit" class="button">Predict</button>

</form>

</div>

</body>

</html>

**HOME.HTML**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<title>Air Quality Index</title>

<link rel="stylesheet" href="{{ url\_for('static', filename='styles2.css') }}">

</head>

<body>

<header>

<div class="header-container">

<h1>Air Quality Index Analyzer</h1>

<nav>

<ul>

<li><a href="{{ url\_for('index') }}">Home</a></li>

<li><a href="{{ url\_for('predict') }}">Predict</a></li>

</ul>

</nav>

</div>

</header>

<div class="container">

<div class="content">

<div class="content-left">

<img src="{{ url\_for('static', filename='AQI.jpg') }}" alt="Air Quality Image" class="aqi-image">

</div>

<div class="content-right">

<h2>Introduction</h2>

<h1>What is AQI?</h1>

<p>The Air Quality Index (AQI) is a measure used to communicate how polluted the air currently is or how polluted it is forecast to become. The AQI is calculated based on the levels of various pollutants in the air, including particulate matter (PM10 and PM2.5), ground-level ozone (O3), carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2).</p>

<h3>Why is AQI important?</h3>

<p>Monitoring and understanding the AQI is crucial for public health as it provides essential information on the safety and quality of the air we breathe. High levels of air pollution can cause a variety of health problems, particularly for vulnerable groups such as children, the elderly, and those with respiratory or cardiovascular conditions. The AQI helps individuals make informed decisions about outdoor activities and take necessary precautions to protect their health.</p>

</div>

</div>

</div>

</body>

</html>

**APP.PY**

from flask import Flask, render\_template, request

import joblib

import pandas as pd

app = Flask(\_\_name\_\_)

# Load model and encoded values

model = joblib.load("model\_small.pkl")

label\_encoder = joblib.load('label\_encoder.pkl')

@app.route('/')

def index():

return render\_template("index.html")

@app.route('/home')

def home():

return render\_template("home.html")

@app.route('/predict')

def predict():

return render\_template("predict.html")

@app.route('/output', methods=["POST"])

def output():

if request.method == 'POST':

city = request.form["city"].strip()

pm25 = float(request.form["pm25"])

pm10 = float(request.form["pm10"])

no2 = float(request.form["no2"])

co = float(request.form["co"])

so2 = float(request.form["so2"])

o3 = float(request.form["o3"])

date = request.form["date"]

# Ensure the city name is valid

if city not in label\_encoder.classes\_:

return render\_template("output.html", y="Invalid City", z="Please enter a valid city name.")

# Transform city and date fields

city\_encoded = label\_encoder.transform([city])[0]

year = int(date.split('-')[0])

month = int(date.split('-')[1])

# Create a DataFrame for the input features

feature\_cols = ['City', 'PM2.5', 'PM10', 'NO2', 'CO', 'SO2', 'O3', 'Year', 'Month']

data = pd.DataFrame([[city\_encoded, pm25, pm10, no2, co, so2, o3, year, month]], columns=feature\_cols)

# Make prediction

pred = model.predict(data)

pred = pred[0]

# Determine AQI category

if pred >= 0 and pred < 50:

res = 'GOOD'

elif pred >= 50 and pred < 100:

res = 'SATISFACTORY'

elif pred >= 100 and pred < 200:

res = 'MODERATELY POLLUTED'

elif pred >= 200 and pred < 300:

res = 'POOR'

elif pred >= 300 and pred < 400:

res = 'VERY POOR'

else:

res = 'SEVERE'

return render\_template("output.html", y=f"AQI: {str(pred)}", z=res)

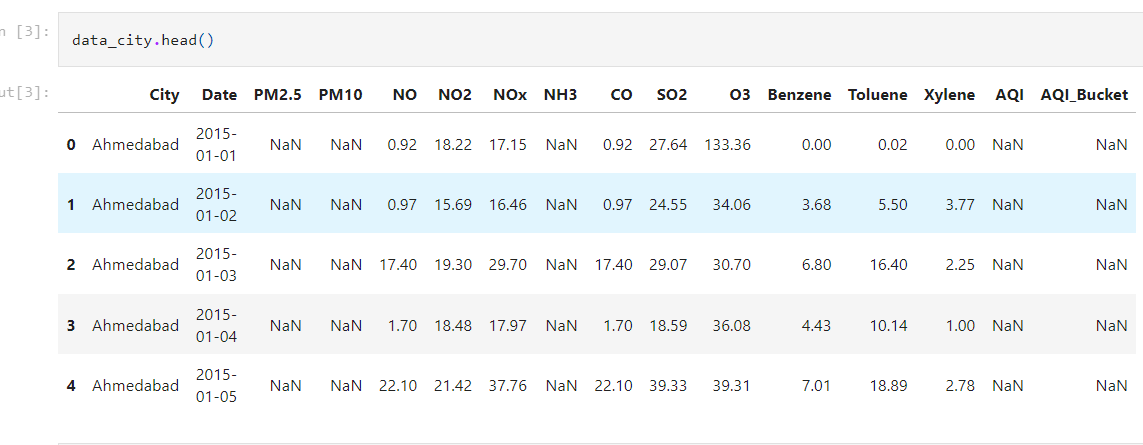
if \_\_name\_\_ == "\_\_main\_\_":

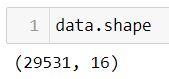
app.run(debug=True)

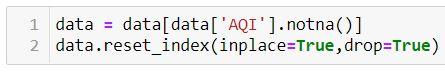
**CODE SNIPPETS**

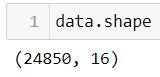
**MODEL BUILDING**

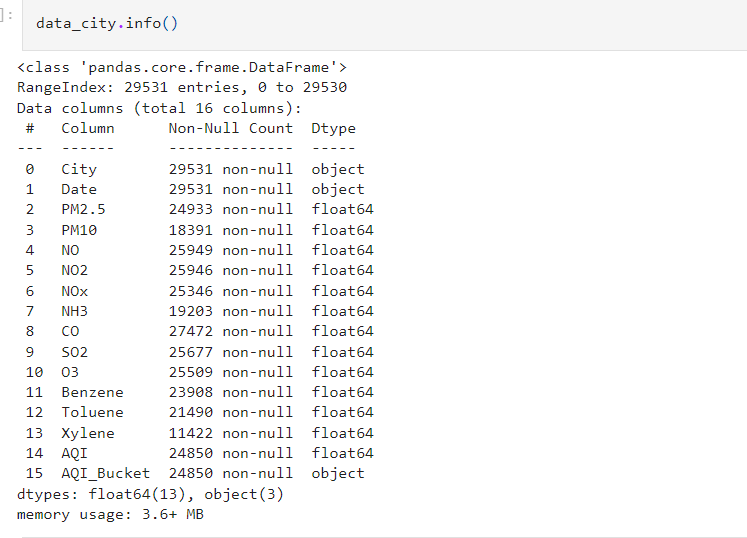




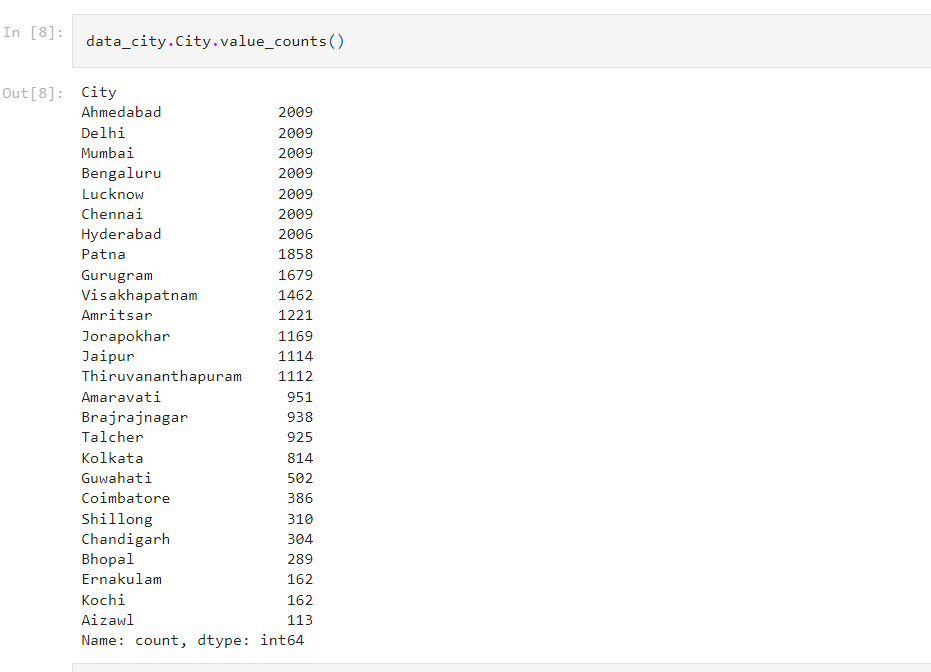












****

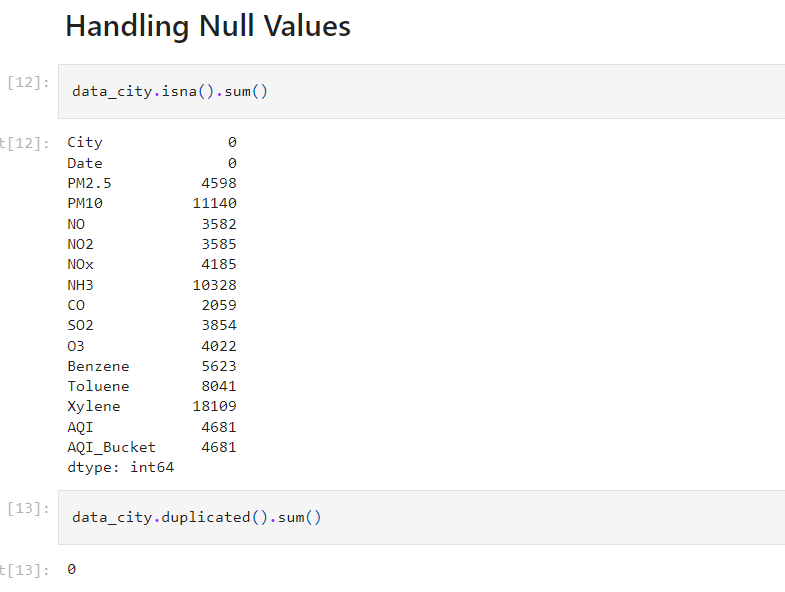
**Data Pre-Processing**

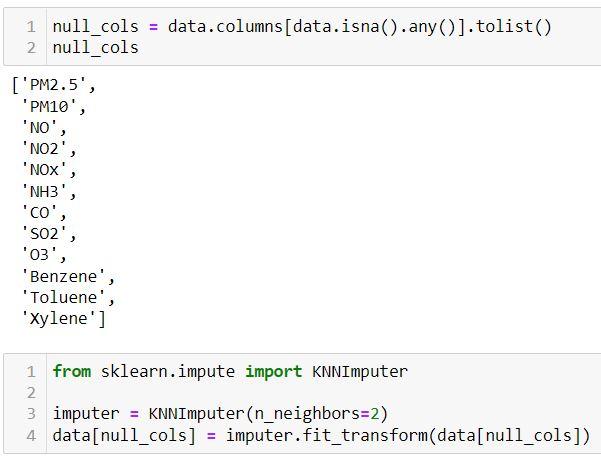
We need to pre-process the collected data before gaining insights and building our model.

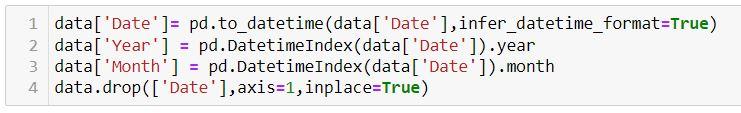
We need to clean the dataset properly in order to fetch good results. This activity includes handling null values and removing unnecessary columns.

**Handling Null Values And Removing Unnecessary Columns**

Let us first see the count of null values in each column:



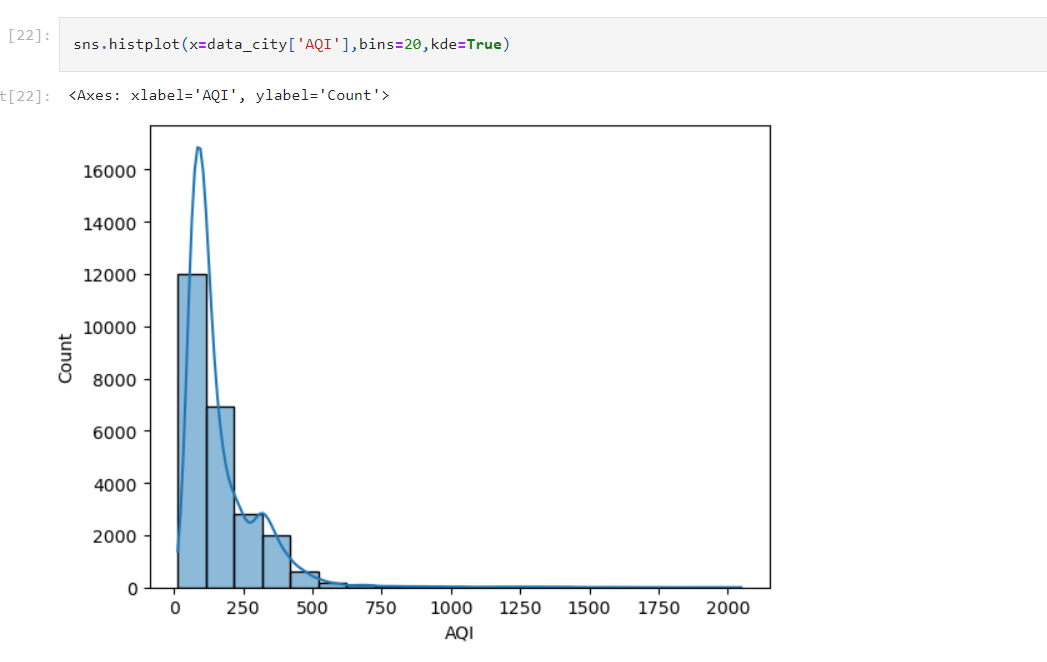


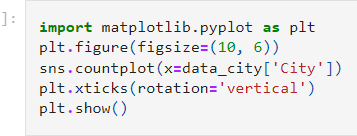




**Data Analysis And Visualization**

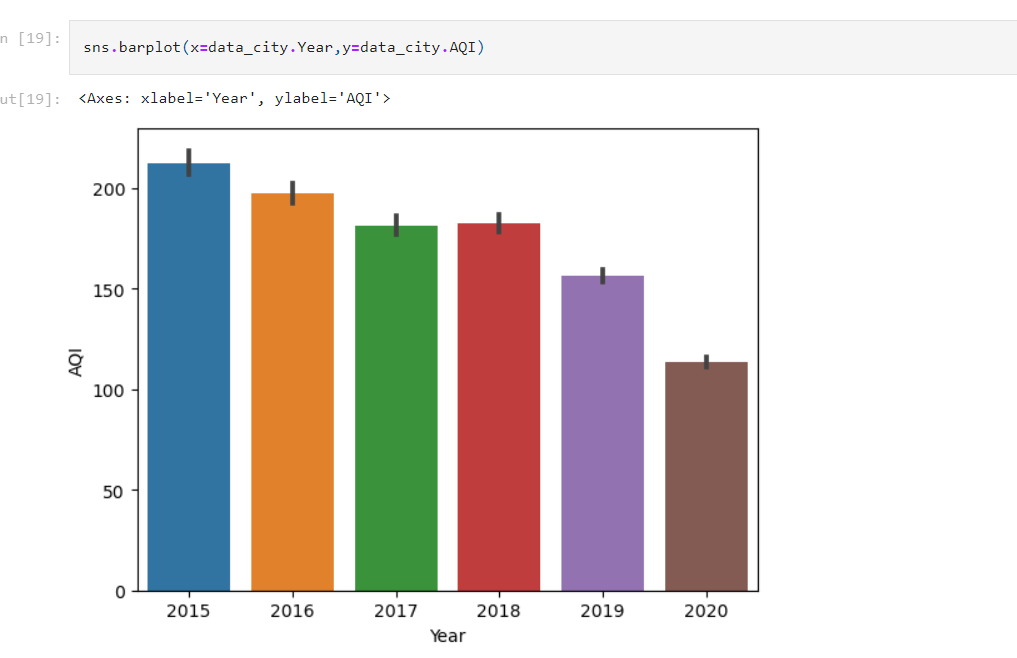
**Univariate Analysis**



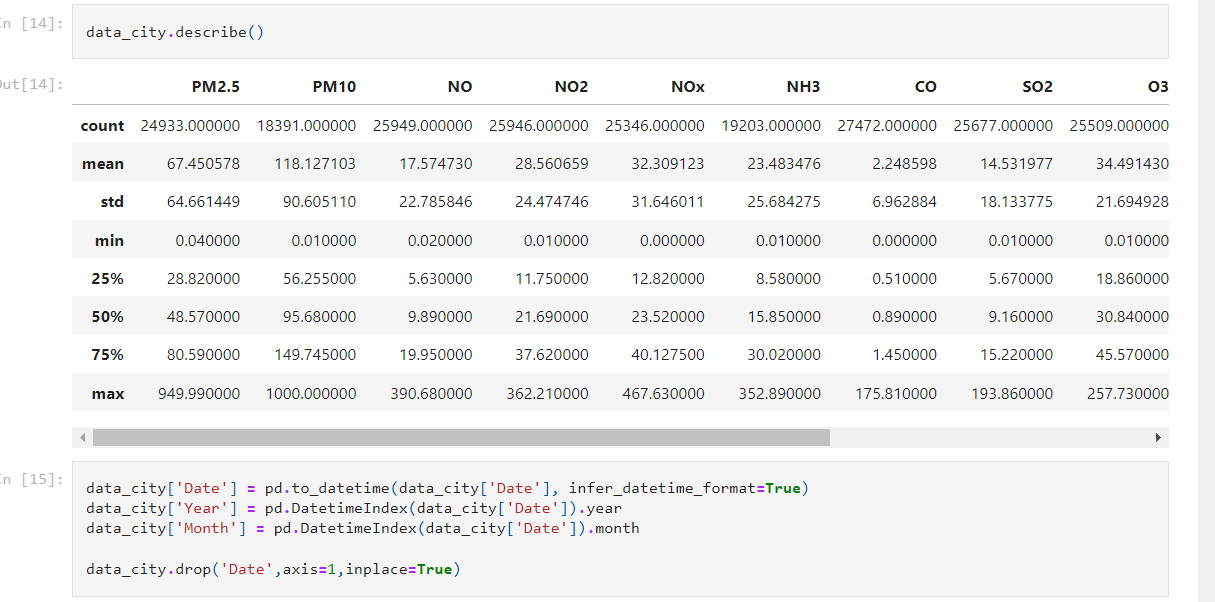


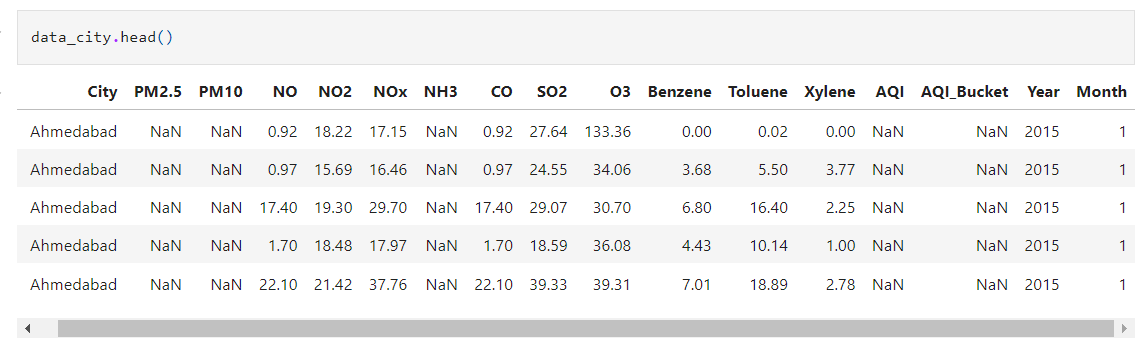


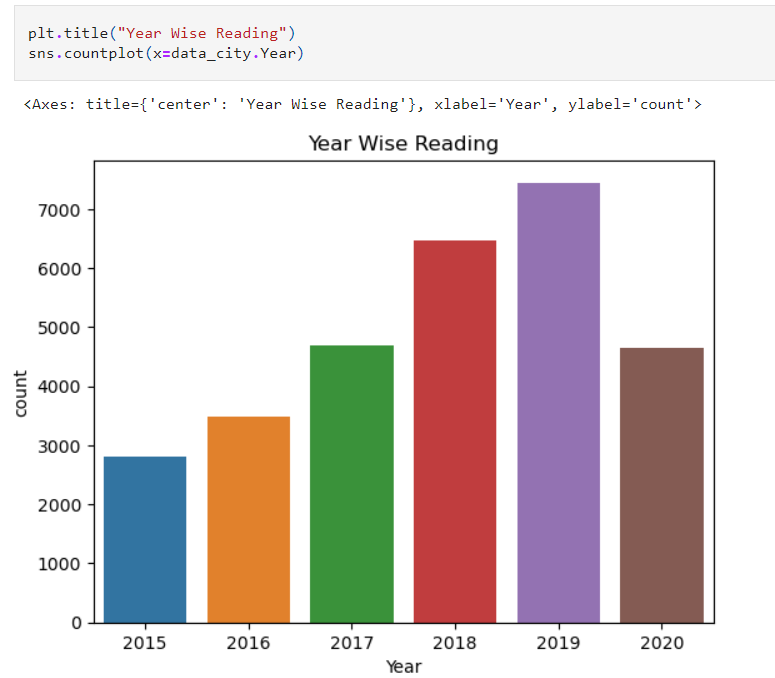
**Bivariate Analysis**

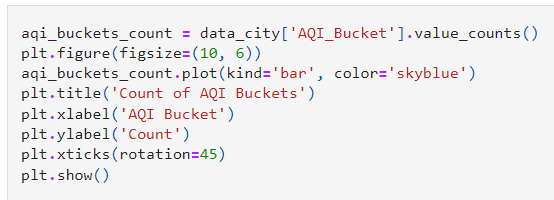


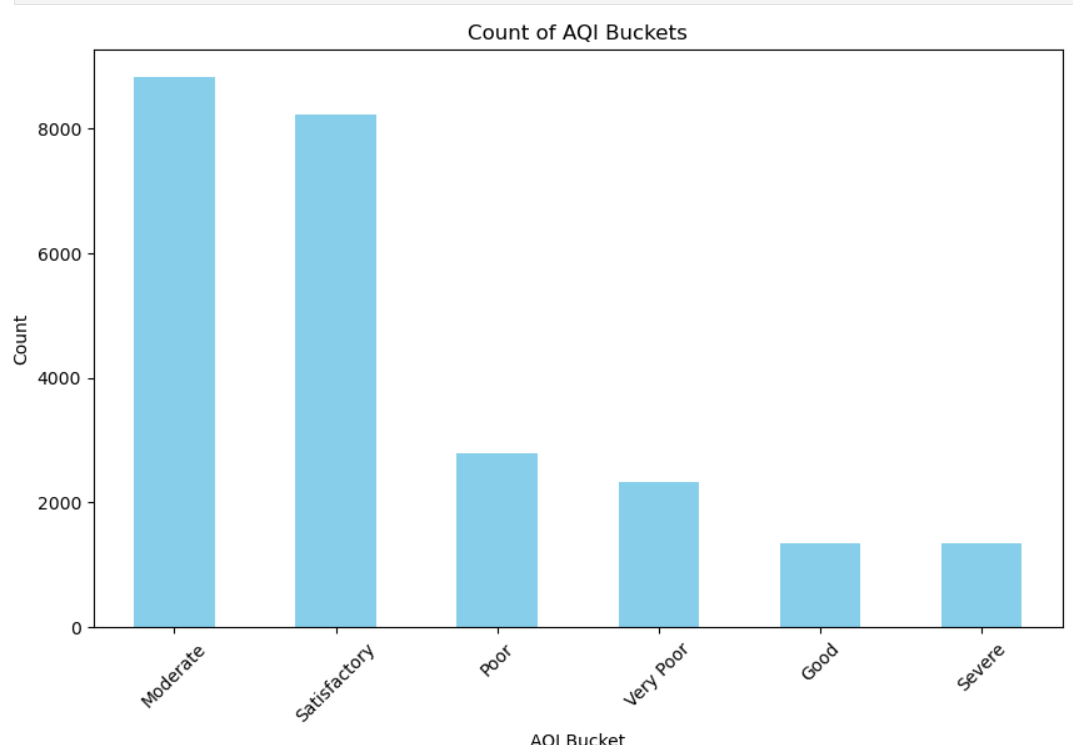
**Descriptive Analysis**

****

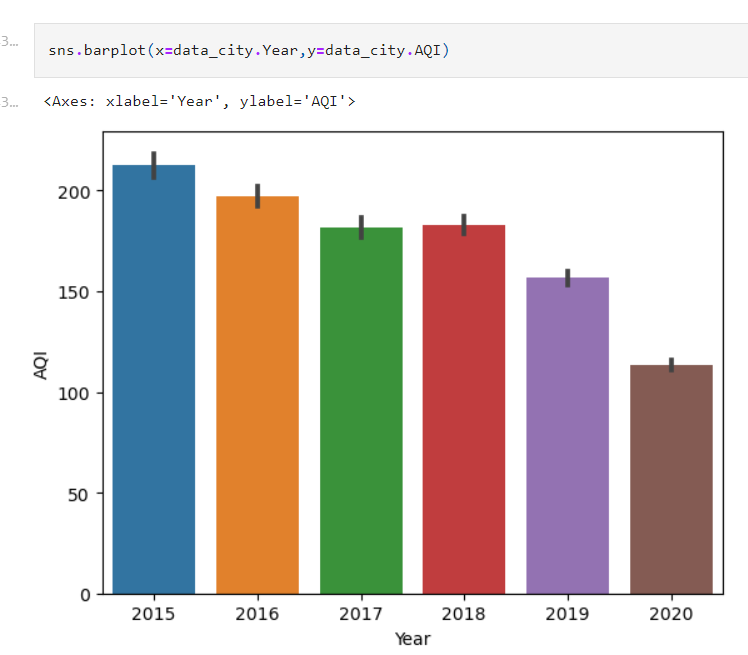
****

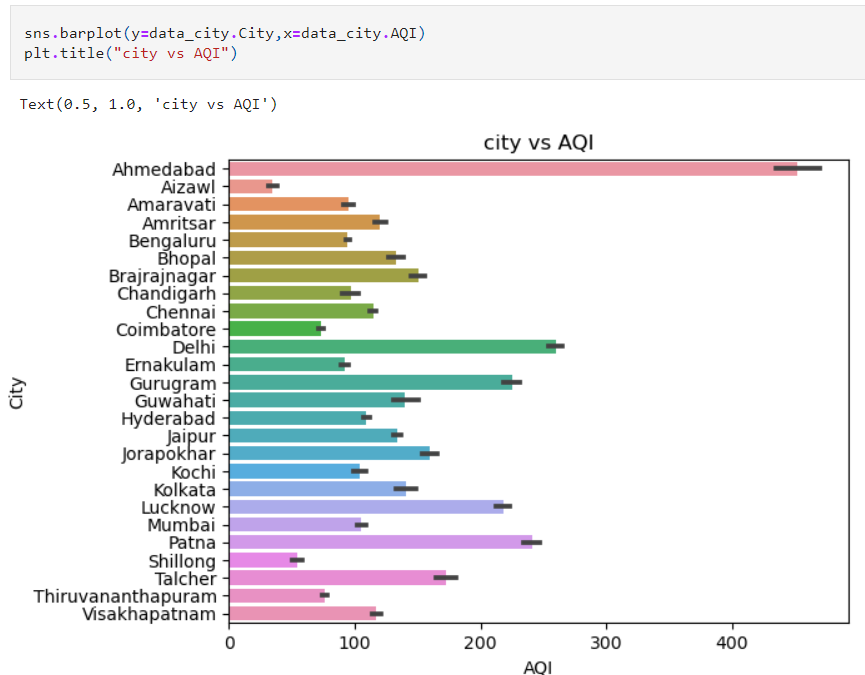
****

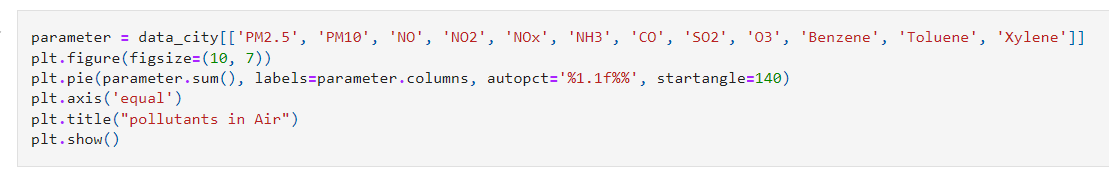
****

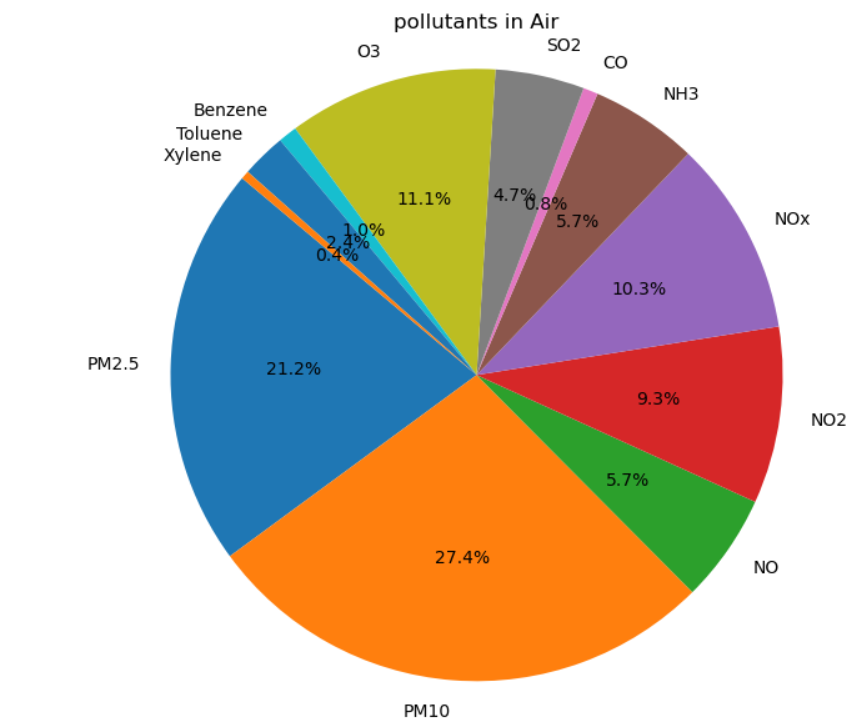
****

**Bar plot**

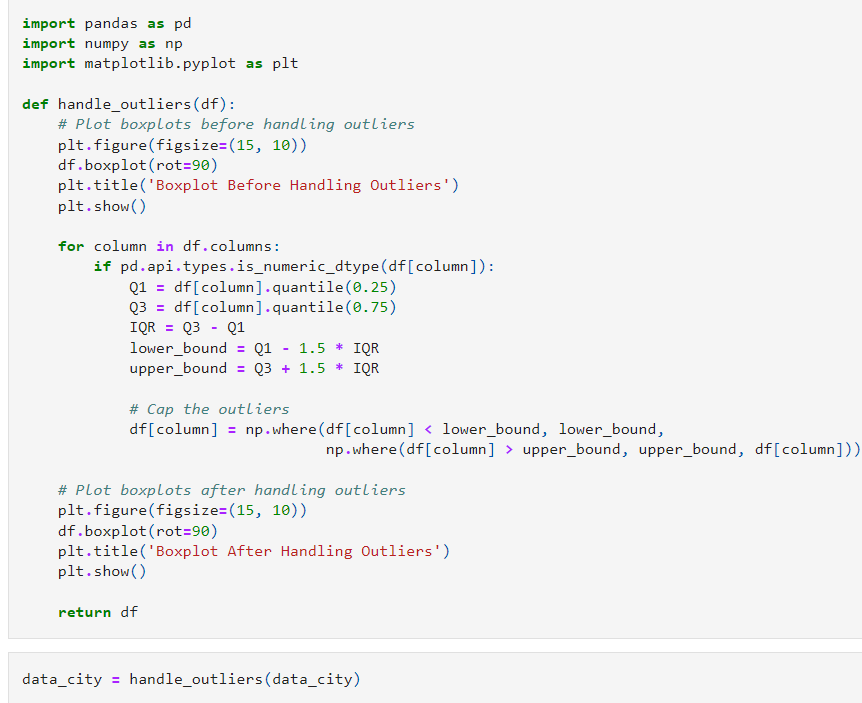
****



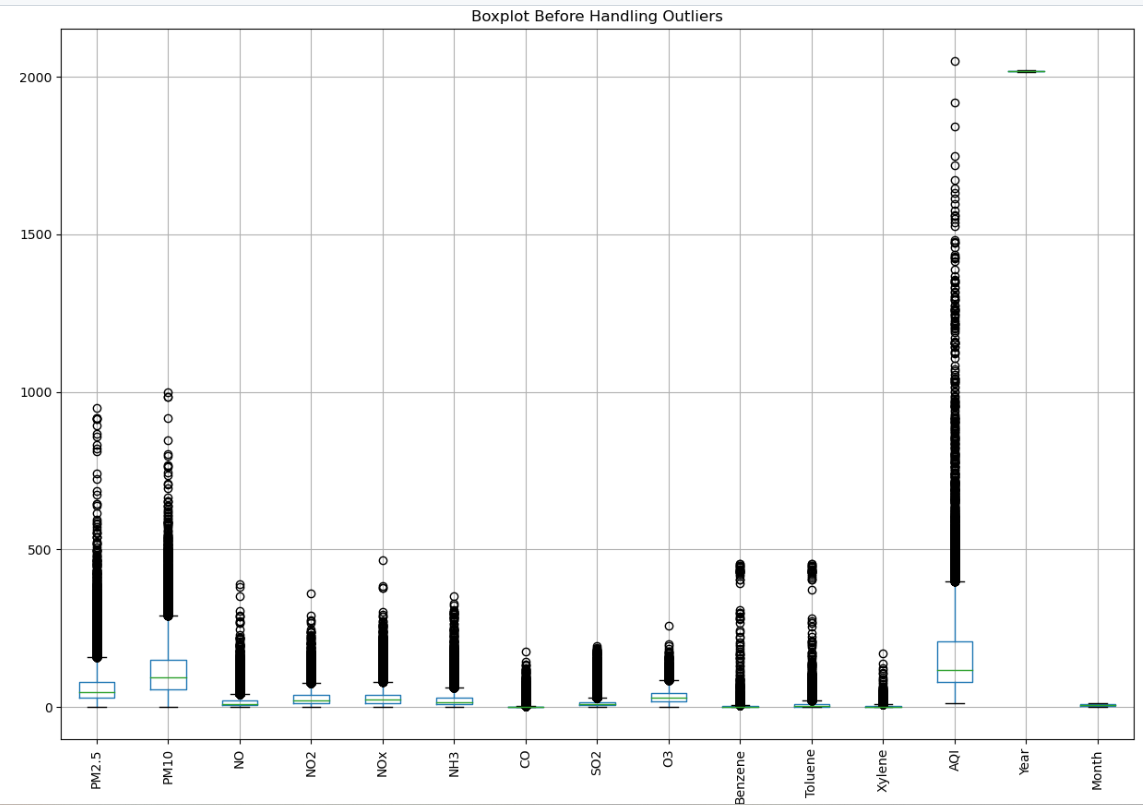




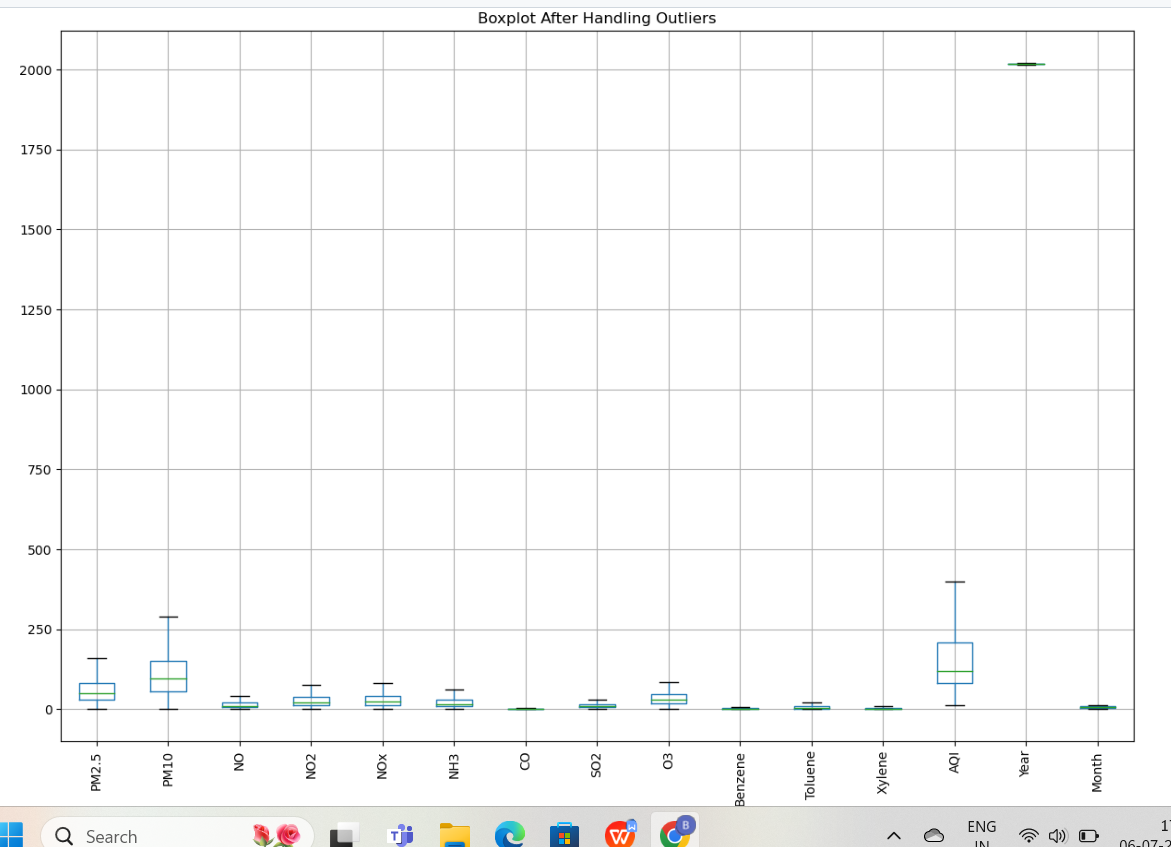
**Box ploting**

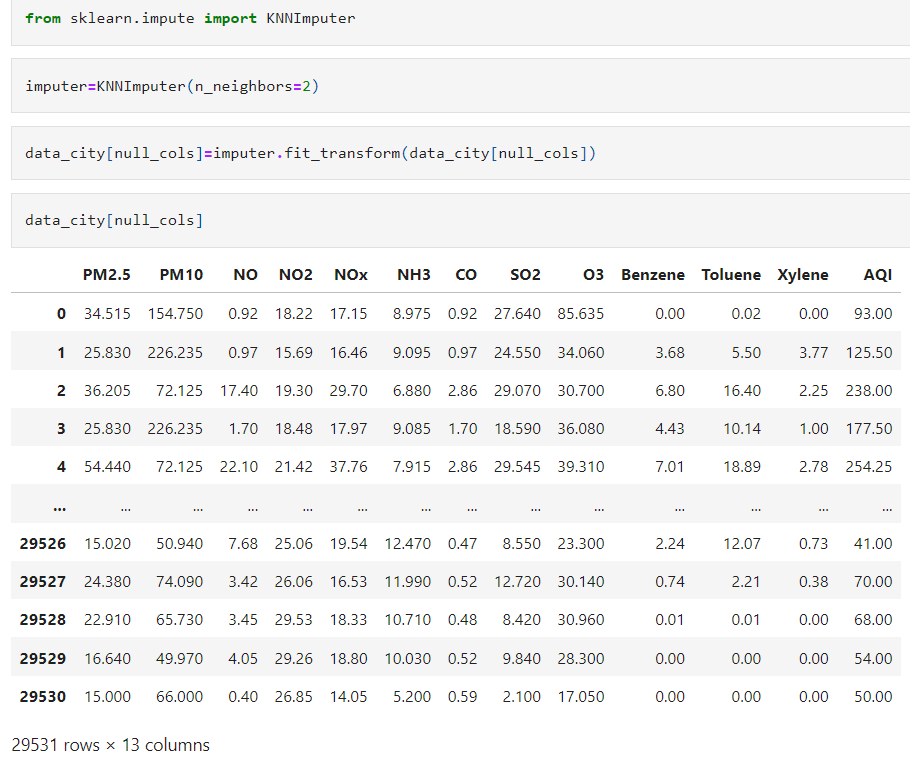


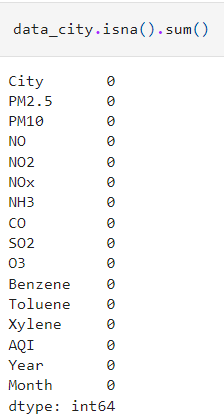
Before Handling Outliers

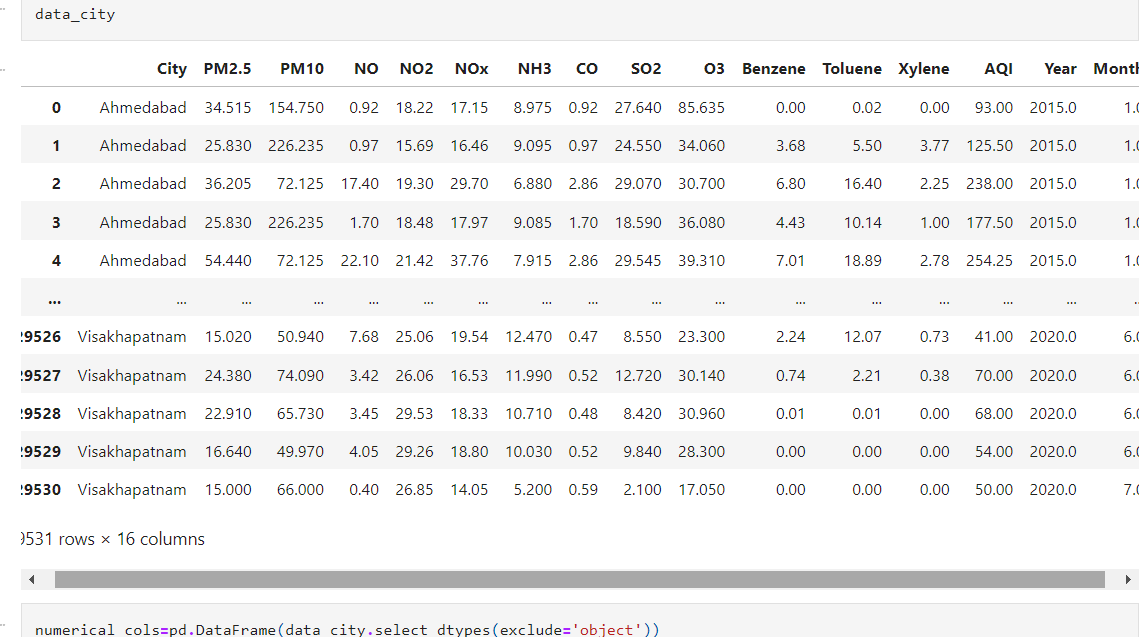


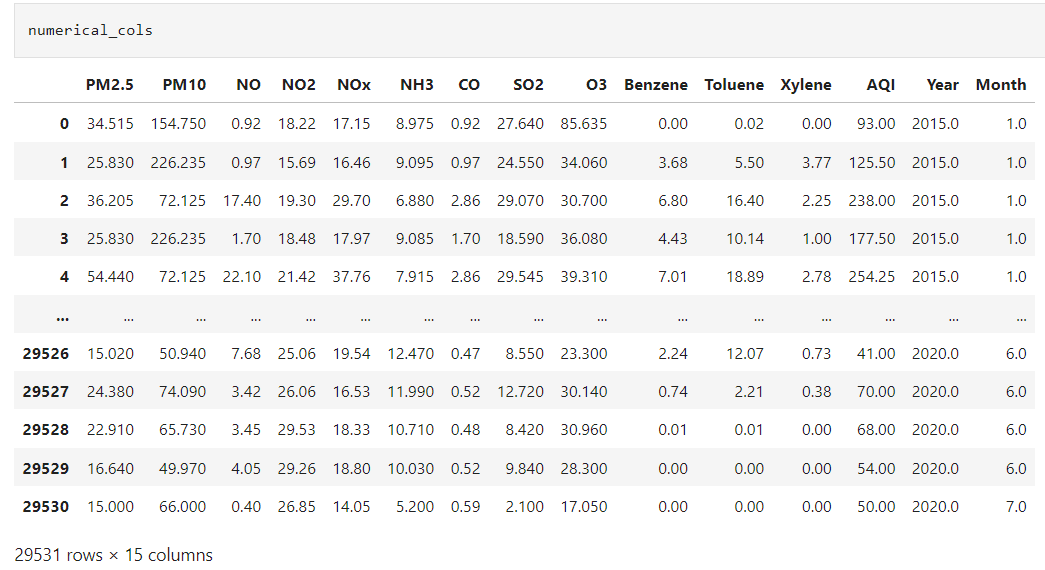
After Handling Outliers

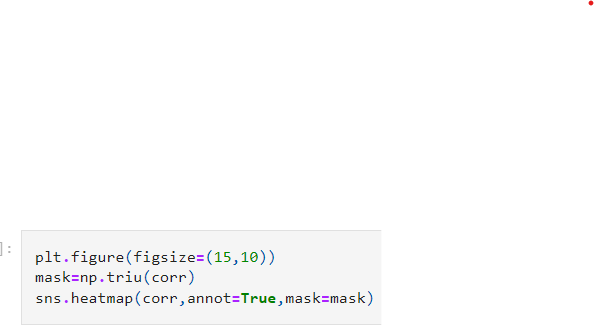


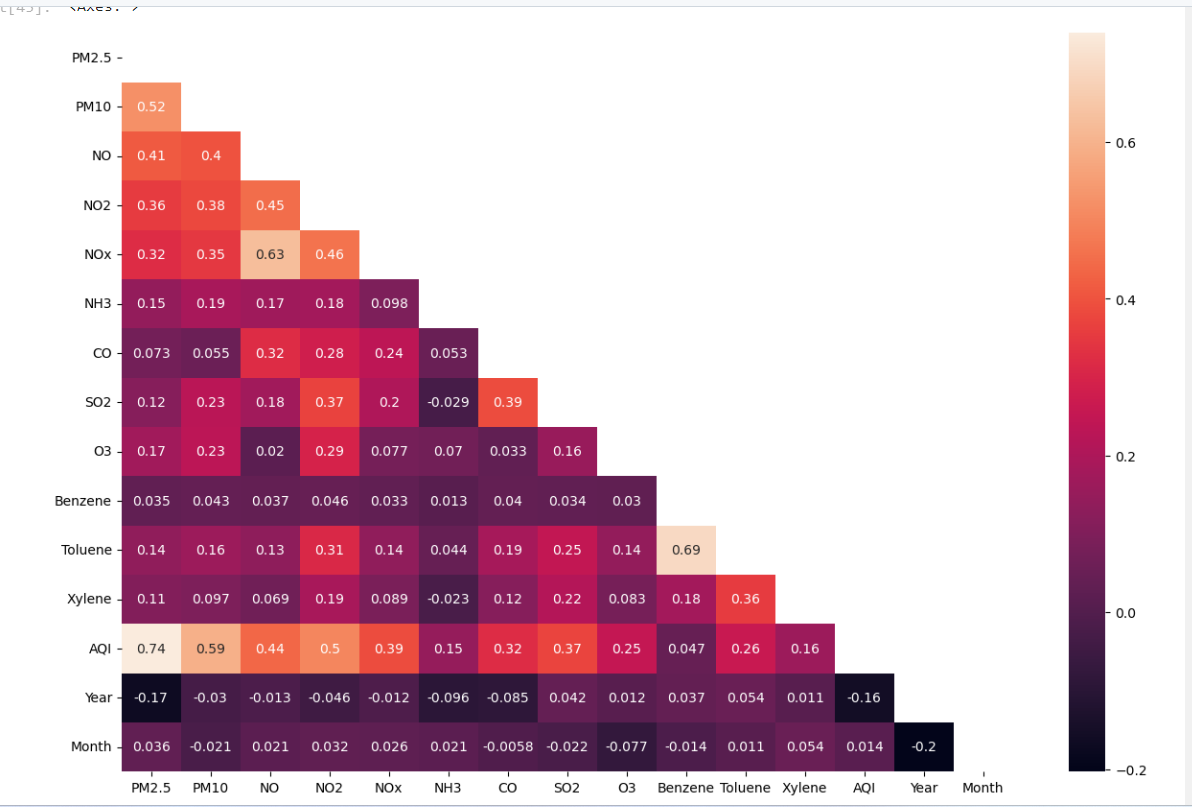


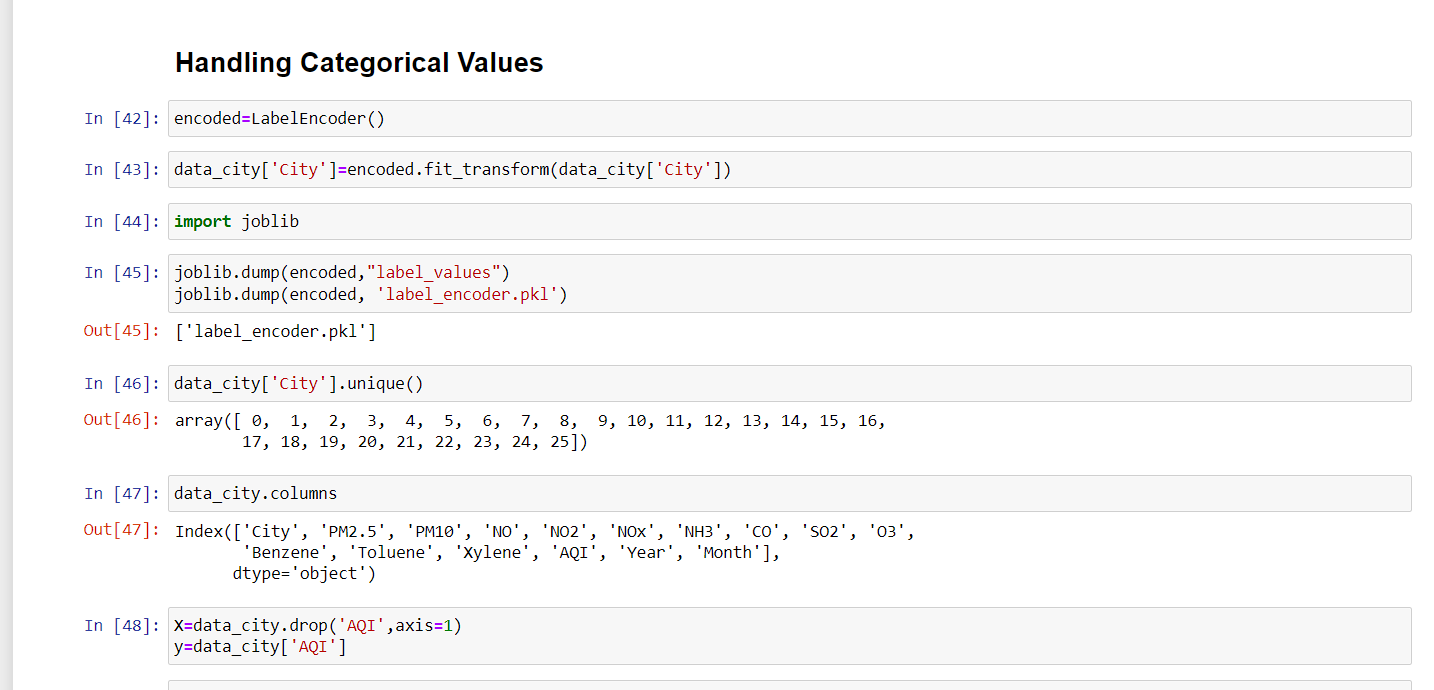


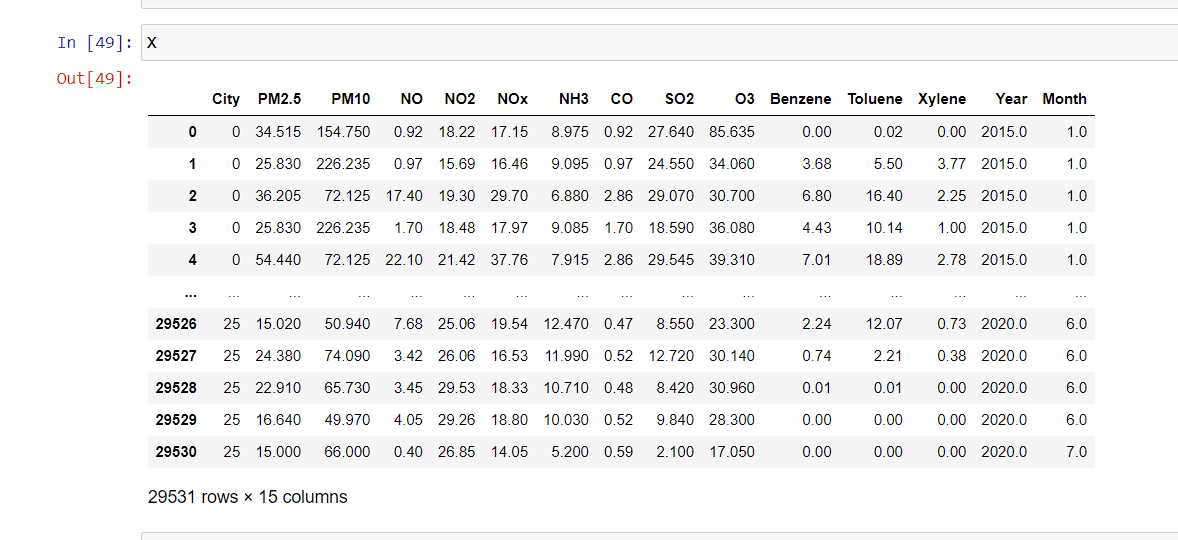


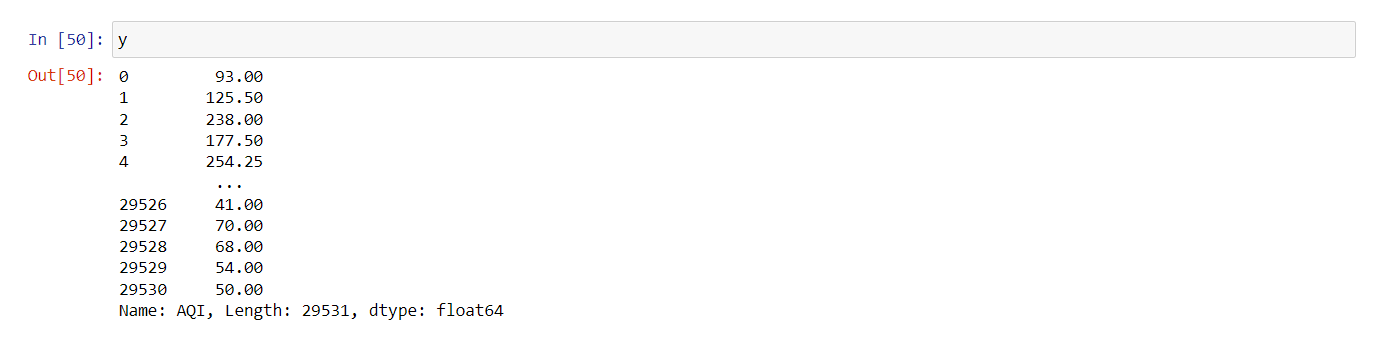


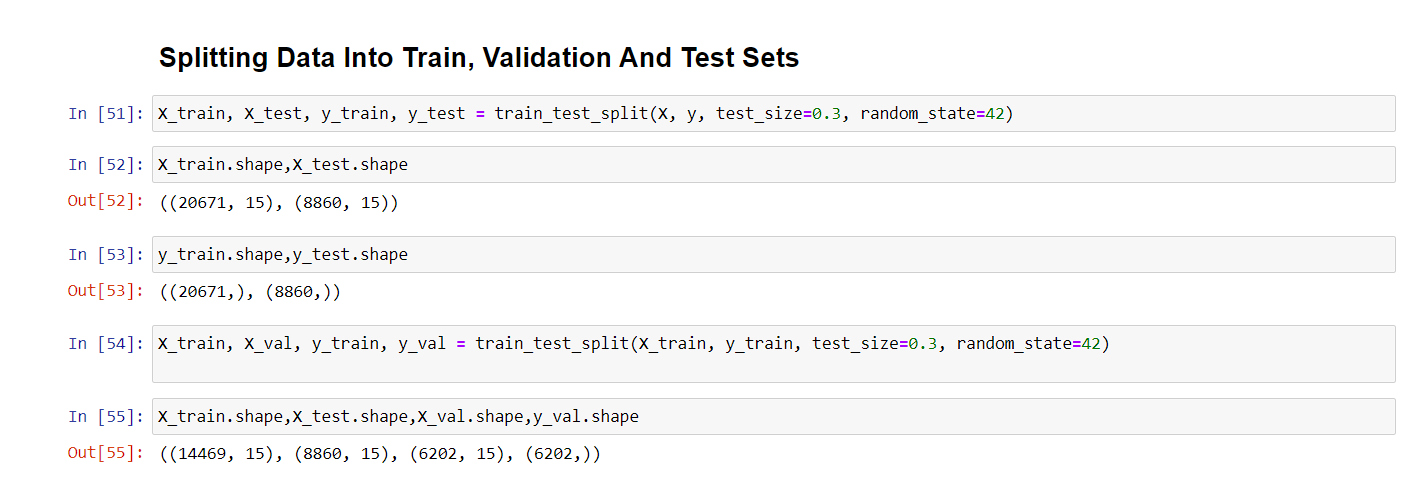


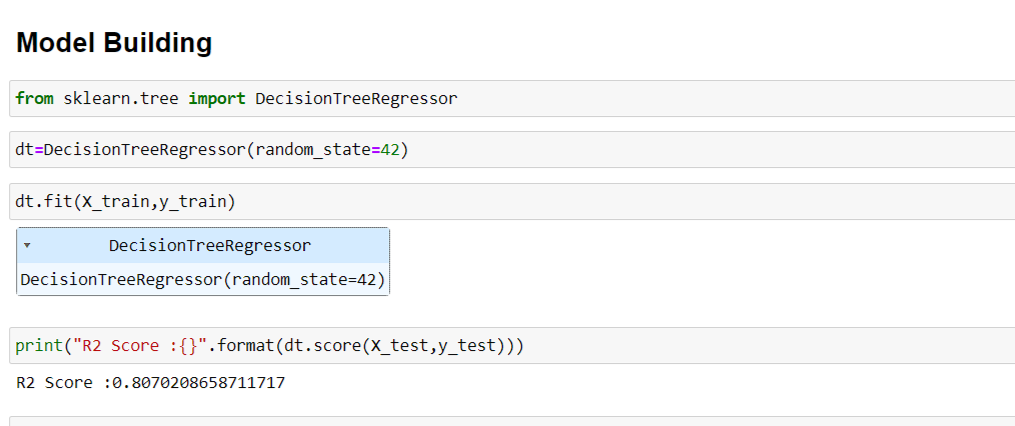


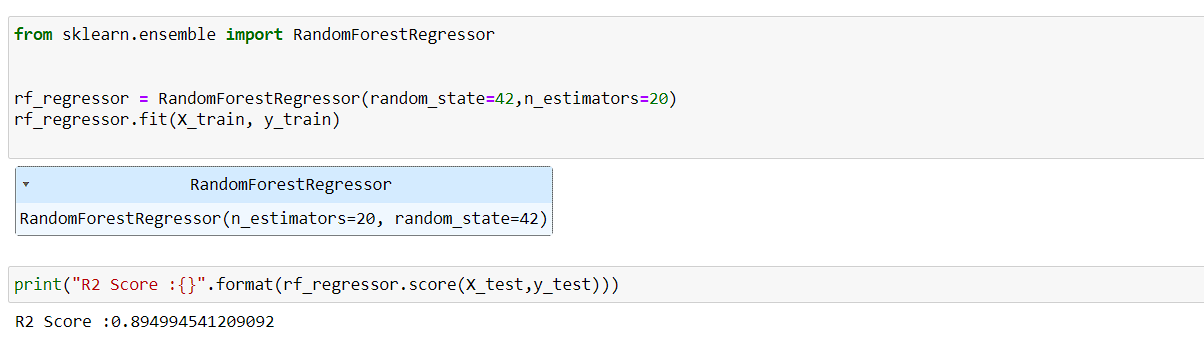




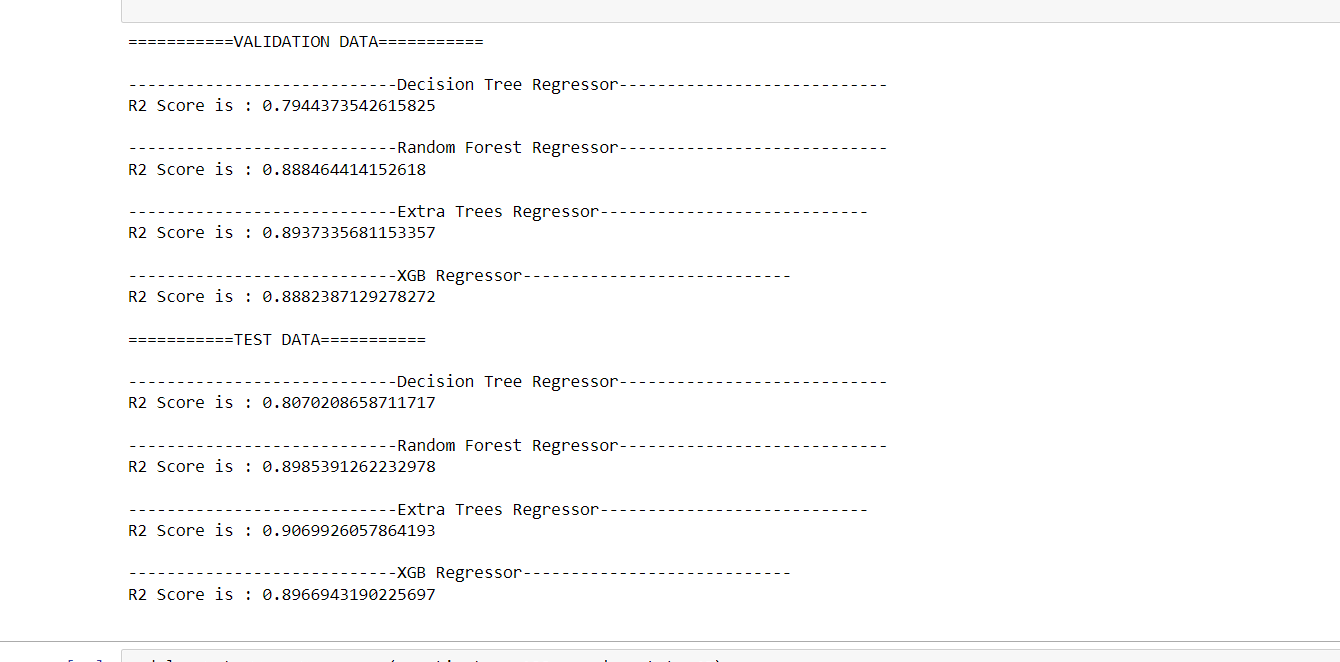
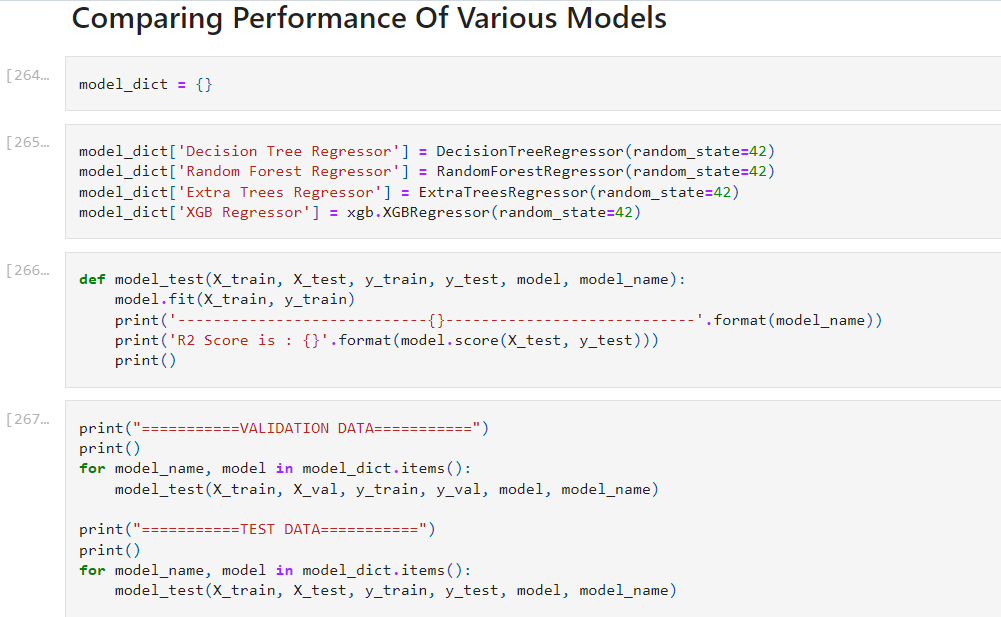
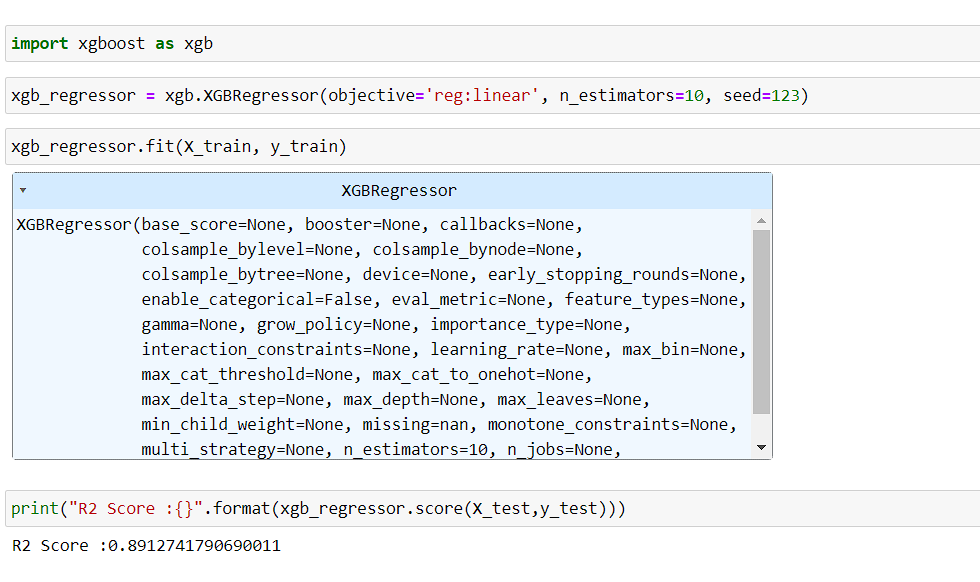


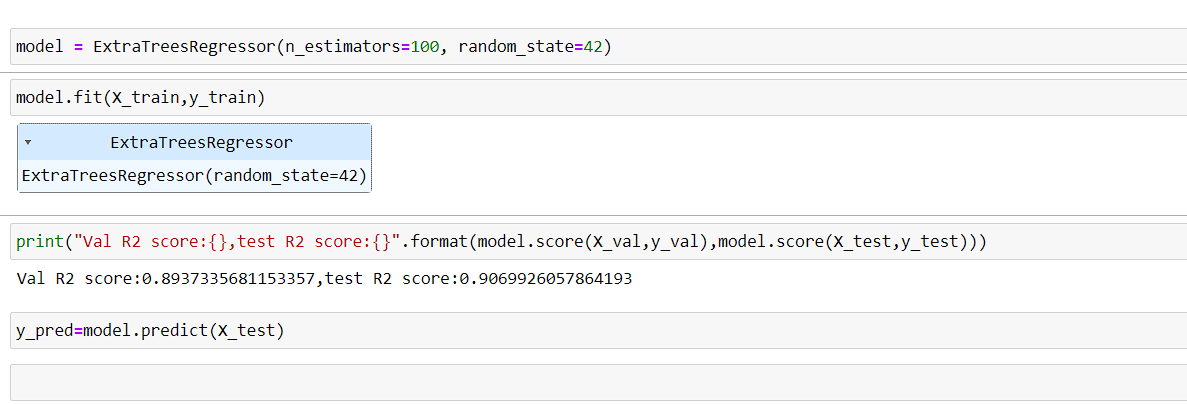












**Evaluating Model Performance**

