

AISSCE-2024-25
PROJECT REPORT

On

Pollution due to Vehicles (2-Wheelers and
4-Wheelers) in Uttar Pradesh over the Last 5
Years

INFORMATICS PRACTICES (065)

SUBMITTED BY:

1. Roll no.:

2. Roll no.:

CLASS:

SUBMITTED TO

TABLE OF CONTENTS

SR. NO	CONTENT	PAGE NO.
1.	INTRODUCTION	5
2.	OBJECTIVE OF THE PROJECT	6
3.	THEORITICAL BACKGROUND	7
4.	SYSTEM IMPLEMENTATION	11
5.	SYSTEM DESIGN & DEVELOPMENT	12
6.	BIBLIOGRAPHY	15

1. INTRODUCTION:

This proposed project “**Pollution due to Vehicles (2-Wheelers and 4-Wheelers) in Uttar Pradesh over the Last 5 Years**” being a small software application developed in Python using Pandas DataFrame and Matplotlib Library as a data analysis and visualization of real-life problem learnt as a part of IP Subject in class XII.

The aim of this project is to analyze the trends in pollution caused by vehicles (2-wheelers and 4-wheelers) over the past 5 years in Uttar Pradesh (UP) using Python. The project includes data analysis, visualization through bar graphs, and insights derived from the analysis.

The purpose of developing this software project is to apply knowledge and understanding of the algorithm design and programming skill in real application.

2. OBJECTIVES OF THE PROJECT

The objective of this project is to apply the programming knowledge into a real- world situation/problem while collecting, classification, and presentation of data in developing a good software to solve real-life applications.

1. Write programs utilizing Python IDLE & software tools.
2. Apply programming principles effectively when developing small to medium sized projects.
3. Write effective procedural code to solve small to medium sized problems.
4. To demonstrate a breadth of knowledge of various Data Analysis and data visualization libraries of Python.
5. To demonstrate ability to conduct a research or applied Computer application project, requiring writing and presentation skills which exemplify scholarly style in IT applications.

3. THEORITICAL BACKGROUND:

About Python Programming Language:



Python is a dynamic, high level, free open source and interpreted programming language. It supports object-oriented programming as well as procedural oriented programming.

There are many features in Python, some of which are discussed below –

1. Easy to code:

Python is high level programming language. Python is very easy to learn language as compared to other language like c, c#, java script, java etc. It is very easy to code in python language and anybody can learn python basic in few hours or days. It is also developer-friendly language.

2. Free and Open Source:

Python language is freely available at official website and you can download it from the given download link below click on the Download Python keyword.

3. Object-Oriented Language:

One of the key features of python is Object-Oriented programming. Python supports object oriented language and concepts of classes, objects encapsulation etc.

4. GUI Programming Support:

Graphical Users interfaces can be made using a module such as PyQt5, PyQt4, wxPython or Tk in python.

5. High-Level Language:

Python is a high-level language. When we write programs in python, we do not need to remember the system architecture, nor do we need to manage the memory.

6. Extensible feature:

Python is a Extensible language. we can write our some python code into c or c++ language and also we can compile that code in c/c++ language.

7. Python is Portable language:

Python language is also a portable language. for example, if we have python code for windows and if we want to run this code on other platform such as Linux, Unix and Mac then we do not need to change it, we can run this code on any platform.

8. Python is Integrated language:

Python is also an Integrated language because we can easily integrated python with other language like c, c++ etc.

9. Interpreted Language:

Python is an Interpreted Language. because python code is executed line by line at a time. like other language c, c++, java etc there is no need to compile python code this makes it easier to debug our code. The source code of python is converted into an immediate form called bytecode.

10. Large Standard Library:

Python has a large standard library which provides rich set of module and functions so you do not have to write your own code for every single thing. There are many libraries present in python for such as regular expressions, unit-testing, web browsers etc.

11. Dynamically Typed Language:

Python is dynamically-typed language. That means the type (for example- int, double, long etc) for a variable is decided at run time not in advance. because of this feature we don't need to specify the type of variable.



Data Analysis and Pandas:

Pandas is an open-source, high-level data analysis and manipulation library for Python programming language. With pandas, it is effortless to load, prepare, manipulate, and analyze data. It is one of the most preferred and widely used libraries for data analysis operations. Pandas have easy syntax and fast operations. Some features of Pandas library are-

- A fast and efficient **DataFrame** object for data manipulation with integrated indexing;
- Tools for **reading and writing data** between in-memory data structures and different formats: CSV and text files, Microsoft Excel, SQL databases, and the fast HDF5 format;
- Intelligent **data alignment** and integrated handling of **missing data**: gain automatic label-based alignment in computations and easily manipulate messy data into an orderly form;
- Flexible **reshaping** and pivoting of data sets;
- Intelligent label-based **slicing, fancy indexing**, and **subsetting** of large data sets;
- Columns can be inserted and deleted from data structures for **size mutability**;
- Aggregating or transforming data with a powerful **group by** engine allowing split-apply-combine operations on data sets;
- High performance **merging and joining** of data sets;
- **Hierarchical axis indexing** provides an intuitive way of working with high-dimensional data in a lower-dimensional data structure;
- **Time series**-functionality: date range generation and frequency conversion, moving window statistics, date shifting and lagging. Even create domain-specific time offsets and join time series without losing data;
- Highly **optimized for performance**, with critical code paths written in Cython or C.

- Python with *pandas* is in use in a wide variety of **academic and commercial** domains, including Finance, Neuroscience, Economics, Statistics, Advertising, Web Analytics, and more.



Data Visualization & Matplotlib:

Matplotlib is a powerful plotting library in Python used for creating static, animated, and interactive visualizations. Matplotlib's primary purpose is to provide users with the tools and functionality to represent data graphically, making it easier to analyze and understand. It was originally developed by John D. Hunter in 2003 and is now maintained by a large community of developers.

Key Features of Matplotlib:

- Versatility:** Matplotlib can generate a wide range of plots, including line plots, scatter plots, bar plots, histograms, pie charts, and more.
- Customization:** It offers extensive customization options to control every aspect of the plot, such as line styles, colors, markers, labels, and annotations.
- Integration with NumPy:** Matplotlib integrates seamlessly with NumPy, making it easy to plot data arrays directly.
- Publication Quality:** Matplotlib produces high-quality plots suitable for publication with fine-grained control over the plot aesthetics.
- Extensible:** Matplotlib is highly extensible, with a large ecosystem of add-on toolkits and extensions like Seaborn, Pandas plotting functions, and Basemap for geographical plotting.

6. **Cross-Platform:** It is platform-independent and can run on various operating systems, including Windows, macOS, and Linux.
7. **Interactive Plots:** Matplotlib supports interactive plotting through the use of widgets and event handling, enabling users to explore data dynamically.

4. SYSTEM IMPLEMENTATION:

Hardware used:

While developing the software, the used hardware's are:

PC with Intel Core i3 processor having 4.00 GB RAM and other required devices.

Software used:

- Microsoft Windows® 10 as Operating System.
- Python IDLE as Front-end Development environment.
- **Pandas:** To create and manipulate dataframes.
- **Matplotlib:** To generate bar graphs.
- **CSV File:** Source data file containing pollution statistics.
- MS-Word 2010 for documentation.

5. SYSTEM DESIGN & DEVELOPMENT:

5.1 Methodology Used:

- **Data Collection:** Data regarding pollution caused by 2-wheelers and 4-wheelers over the last 5 years in Uttar Pradesh is stored in a CSV file.
- **Data Analysis:** Use Pandas to load and analyze the data.
- **Visualization:** Create bar graphs using Matplotlib to visualize trends and derive insights.
- **Result Interpretation:** Analyze the graphical trends to understand the impact of vehicle emissions.

5.2 Coding:

```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt

# Load the CSV data into a Pandas dataframe
data = pd.read_csv("pollution_data.csv")

# Display the first few rows of the dataframe
print("Data Preview:")
print(data.head())

# Data Analysis: Group data by year and vehicle type
pollution_by_year = data.groupby(["Year",
                                   "Vehicle_Type"])["Pollution_Level"].sum().unstack()

# Plotting the data
plt.figure(figsize=(10, 6))
bar_width = 0.4

# Bar graph for 2-Wheelers
pollution_by_year['2-Wheelers'].plot(kind='bar',
width=bar_width, label='2-Wheelers', color='blue', position=0)

# Bar graph for 4-Wheelers
pollution_by_year['4-Wheelers'].plot(kind='bar',
width=bar_width, label='4-Wheelers', color='green', position=1)

# Adding labels, title, and legend
plt.title("Pollution Levels by Vehicles in Uttar Pradesh (Last 5 Years)")
plt.xlabel("Year")
plt.ylabel("Pollution Level (in units)")
plt.xticks(rotation=45)
plt.legend()
plt.tight_layout()
```

```
# Save the graph as an image file
plt.savefig("pollution_graph.png")

# Display the graph
plt.show()
```

5.3 Sample Dataset (Input):

Sample CSV Data (pollution_data.csv)

Year	Vehicle_Type	Pollution_Level
2019	2-Wheelers	500
2019	4-Wheelers	700
2020	2-Wheelers	550
2020	4-Wheelers	750
2021	2-Wheelers	600
2021	4-Wheelers	800
2022	2-Wheelers	650
2022	4-Wheelers	850
2023	2-Wheelers	700
2023	4-Wheelers	900

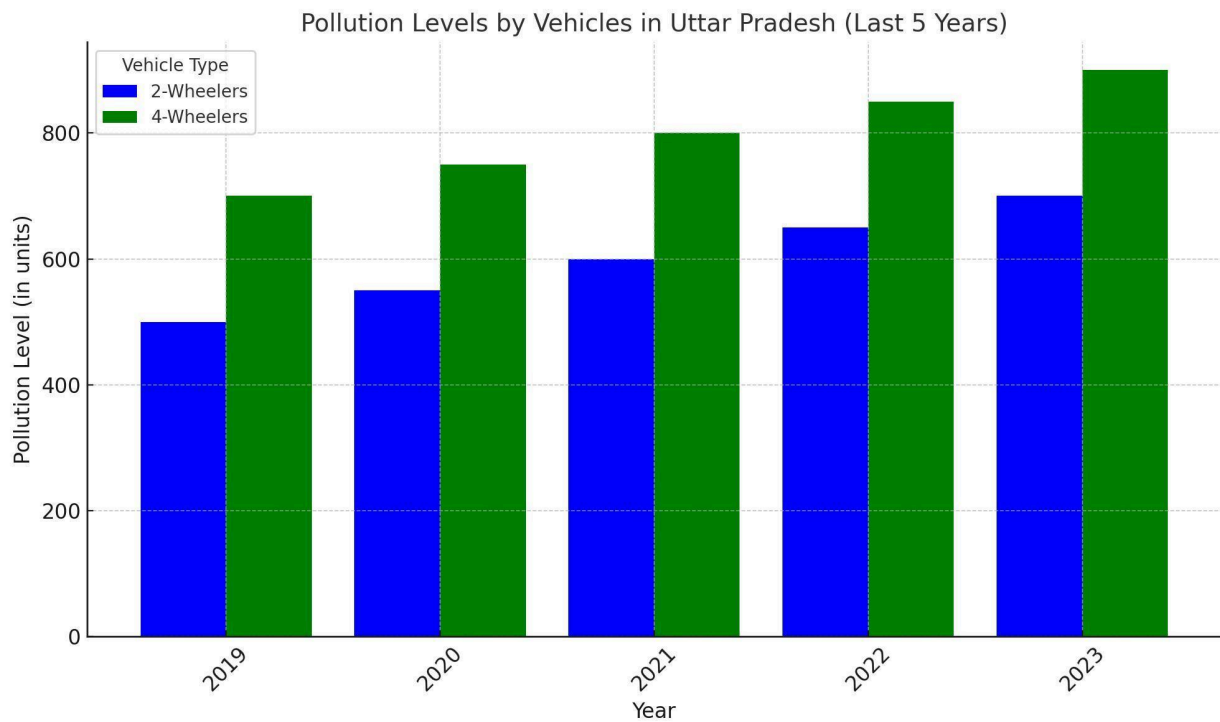
Dataframe Preview:

```

      Year  Vehicle_Type  Pollution_Level
0  2019    2-Wheelers         500
1  2019    4-Wheelers         700
2  2020    2-Wheelers         550
3  2020    4-Wheelers         750
4  2021    2-Wheelers         600

```

5.4 Outputs:



5.5 Observations:

1. Pollution levels from both 2-wheelers and 4-wheelers have shown a steady increase over the past 5 years.
2. 4-wheelers consistently contribute more to pollution compared to 2-wheelers.
3. The rise in pollution may be attributed to the increasing number of vehicles and urbanization in Uttar Pradesh.

5.6. Conclusion

The analysis highlights the growing concern of vehicle emissions contributing to air pollution in Uttar Pradesh. Policies promoting electric vehicles and stricter emission norms could help mitigate the problem.

6. BIBLIOGRAPHY

- 1. Informatics Practices -Class XII- NCERT, Publication*
- 2. Various websites and blogs related to Python Programming.*
- 3. Website: <https://www.python.org>*
- 4. Pandas Documentation & Matplotlib Documentation*