Internship Final Report

Student Name: Ayushi Dwivedi

University: Shri Ramswaroop Memorial College Of Management

Major: Computer Science

Internship duration: 1 June 2025 to 31 June 2025

Company: Shadowfox

Domain: Data Science

Mentor: Mr. Hari

Coordinator: Mr. Aakash

Objectives

To gain practical experience and enhance my analytical and technical skills in the field of Data Science by applying theoretical knowledge to real-world data problems. The objective of this internship was to develop a strong foundation in data handling, analysis, and visualization while contributing meaningfully to organizational projects. This opportunity also aimed to improve my proficiency in data science tools and techniques, teamwork, and problem-solving abilities.

Tasks and Responsibilities

Project Planning and Research Definition

Identified core environmental challenges related to air pollution in Delhi.

Formulated research questions focusing on major pollutants (PM2.5, PM10, NO_2, etc.), seasonal trends, and geographical influences on AQI levels.

Data Collection and Preprocessing

Sourced AQI datasets from government and open-data portals.

Cleaned and organized data using Python libraries such as Pandas and NumPy to ensure accuracy and consistency.

Statistical and Exploratory Data Analysis

Applied statistical techniques to understand pollutant concentration patterns and seasonal fluctuations.

Conducted correlation analysis between pollutants and environmental/geographical variables.

Data Visualization and Insight Generation

Utilized advanced visualization tools such as Matplotlib, Seaborn, and Plotly to create informative and interactive graphs.

Highlighted temporal and spatial patterns in AQI across different regions and seasons in Delhi.

Documentation of Visualization Libraries

Created comprehensive documentation explaining the usage, customization, and best practices of various Python visualization libraries.

Provided code examples and visual outputs to enhance the learning and usability of the documentation.

Interpretation and Reporting

Interpreted analytical results to offer actionable insights on AQI dynamics.

Compiled findings into a structured report that can support policy recommendations and public health strategies.

Collaboration and Feedback Integration

Coordinated with mentors and team members for regular feedback and project direction.

Incorporated suggestions to refine the analysis and presentation quality.

Learning Outcomes

During my internship at Shadowfox, I gained valuable hands-on experience in data science. I learned how to work with real-world datasets, perform statistical analysis, and extract meaningful insights. I developed skills in using Python libraries like Pandas, NumPy, Matplotlib, Seaborn, and Plotly for data processing and visualization.

I also learned how to document visualization tools effectively and present findings in a clear and structured manner. The project enhanced my problem-solving ability, research mindset, and understanding of air quality issues in Delhi. This internship significantly improved my technical knowledge, communication skills, and overall confidence in handling data-driven projects.

Challenges and Solutions

1. Data Quality and Inconsistency

Challenge: The AQI datasets contained missing values, inconsistent formats, and outliers.

Solution: Used data cleaning techniques with Pandas to handle missing data, remove duplicates, and standardize formats for accurate analysis.

2. Complex Visualizations

Challenge: Representing multi-dimensional data (pollutants, time, geography) in a clear and insightful manner was difficult.

Solution: Explored and documented advanced features of libraries like Seaborn and Plotly to build interactive and informative visualizations.

3. Defining Relevant Research Questions

Challenge: Framing the right questions to drive a meaningful analysis required domain understanding.

Solution: Conducted background research on air pollution in Delhi and collaborated with

mentors to refine research direction.

4. Time Management

Challenge: Balancing project execution with learning new tools and documentation writing was demanding.

Solution: Followed a structured weekly plan, breaking tasks into smaller goals and setting clear deadlines.

Conclusion

The internship at Shadowfox was a valuable learning experience that enhanced both my technical and analytical abilities in the field of Data Science. Working on a real-world project related to the Air Quality Index (AQI) in Delhi allowed me to apply my theoretical knowledge to practical challenges. I gained hands-on experience in data analysis, visualization, and documentation, while also deepening my understanding of environmental data and public health concerns.

This internship not only improved my proficiency in tools like Python, Pandas, and visualization libraries but also strengthened my research, problem-solving, and communication skills. Overall, it was a rewarding experience that has prepared me for future roles in data science and analytics

Acknowledgments

I would like to express my sincere gratitude to Shadowfox for providing me with the opportunity to intern in the Data Science domain. It was a truly enriching experience that allowed me to grow both professionally and personally.

I extend my special thanks to my mentor and the entire data science team at Shadowfox for their constant support, guidance, and constructive feedback throughout the internship. Their insights and encouragement played a crucial role in helping me

successfully complete my project on Air Quality Index analysis.

I am also thankful to my academic faculty and institution for preparing me with the foundational knowledge and for their continuous encouragement. Lastly, I appreciate the support of my family and peers who motivated me throughout this journey.