<DATA SCIENCE TOOLBOX: PYTHON PROGRAMMING>

PROJECT REPORT

(Project Semester January-April 2025)

Electric Vehicle Population

Submitted by

Name: Nithin

Registration No:12304334

Programme and Section: Data Science, K23DP

Course Code: INT375

Under the Guidance of

Assistant Professor.Dr. Dhiraj Kapila(UID:23509)

Discipline of CSE/IT

Lovely School of Computer Science

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Nithin bearing Registration no. 12304334 has completed INT375 project titled, “Electric Vehicle Population” under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort and study.

Signature and Name of the Supervisor
Designation of the Supervisor
School of Computer Science
Lovely Professional University
Phagwara, Punjab.

Date:

DECLARATION

I, Nithin, student of Data Science, under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 12-04-2025 Signature

Registration No. 12304334 Nithin

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project guide, Dhiraj Kapila, for their valuable guidance and support throughout this project, *"Electric vehicle population "*. I am thankful to the Department of Computer Science and Engineering, Lovely Professional University, for providing the necessary resources and environment. I also acknowledge the Government of India for making the dataset publicly available, enabling this research.

Name:Nithin

Registration number: 12304334

Table of Contents

INTRODUCTION

India, with its rapidly growing population and urbanization, is witnessing a significant shift in transportation trends, particularly with the increasing adoption of electric vehicles (EVs). Understanding the distribution and growth of EVs across the country is essential for effective policy formulation, infrastructure development, and environmental planning. The dataset from the year 2013–14 offers a valuable glimpse into early trends and challenges related to the adoption of electric vehicles across Indian states and districts.

This project aims to perform Exploratory Data Analysis (EDA) on the Indian Electric Vehicle Population dataset, focusing on identifying patterns, trends, and anomalies in EV distribution and usage. EDA is a foundational step in any data-driven initiative, enabling insights through data cleaning, transformation, and visualization before proceeding to advanced modeling or prediction.

Using Python libraries such as Pandas, Matplotlib, and NumPy, this analysis reveals meaningful relationships and visual trends that aid in a deeper understanding of the EV landscape in India. The project also sheds light on regional disparities in EV adoption and charging infrastructure, helping identify areas that may need targeted intervention from government bodies or private stakeholders.

The ultimate goal of this project is to extract actionable insights from the data to support informed decision-making, with an emphasis on enhancing sustainable mobility and accelerating India’s transition to a greener transport ecosystem.

Source of Dataset

- File Name: electric vehicle population.csv
- Source: https://catalog.data.gov/dataset/electric-vehicle-population-data

- Attributes:

- State, District
- Number of Registered Electric Vehicles
- Number of Charging Stations
- Lack of Charging Infrastructure – Count
- Battery Degradation Issues – Count

EDA Process (Exploratory Data Analysis Process)

Exploratory Data Analysis (EDA) is a vital step in the data science workflow. It focuses on understanding the structure, trends, and key characteristics of the dataset through statistical summaries and visualizations. EDA helps uncover hidden patterns, detect anomalies, test assumptions, and generate hypotheses, all of which are essential before moving into predictive modeling or deeper analysis.

For this Electric Vehicle Population project, the EDA process followed a structured approach:

3.1. Data Collection

The dataset used in this analysis was compiled from publicly available government sources and EV registration data. It includes information related to the number of registered electric vehicles, availability of charging infrastructure, and challenges affecting EV adoption across various states and districts in India.

3.2. Data Cleaning and Preparation

To ensure data integrity and ease of analysis, the following cleaning and preparation steps were performed:

- Addressed missing or inconsistent values to maintain data quality.
- Standardized column names for uniformity and better readability.
- Filtered key attributes such as ‘State’, ‘District’, ‘Number of EVs’, ‘Charging Stations’, and reported EV challenges.
- Converted data types where necessary to support numerical operations.
- Grouped data by state and district levels for aggregated insights and comparisons.

- **Bar and line charts** to compare EV registration numbers and infrastructure availability across states.
- **Scatter plots** to identify relationships between EV population and charging station availability.ready for meaningful analysis.

3.3. Data Exploration and Visualization

With the data prepared, a variety of visual techniques were applied to uncover insights:

- **Bar and line charts** to compare EV registration numbers and infrastructure availability across states.
- **Scatter plots** to identify relationships between EV population and charging station availability.
- **Histograms** to explore the distribution of reported EV issues, such as battery degradation or lack of infrastructure.
- **Horizontal bar charts** to highlight district-level adoption patterns and problem areas.

These visual tools helped expose regional trends, infrastructure gaps, and correlations among different challenges affecting EV growth.

Overall, the EDA process provided critical insights into the current landscape of electric vehicle adoption in India. It also laid the foundation for future studies or policy development aimed at boosting sustainable transportation through improved planning and infrastructure.

4. ANALYSIS ON DATASET

4.1 Line Chart – Total Electric Vehicle Population per State (Simulated Trend)

4.1.1 Introduction

This analysis visualizes the cumulative number of registered electric vehicles (EVs) per state, helping to identify adoption patterns and growth trends across different regions in India.

4.1.2 General Description

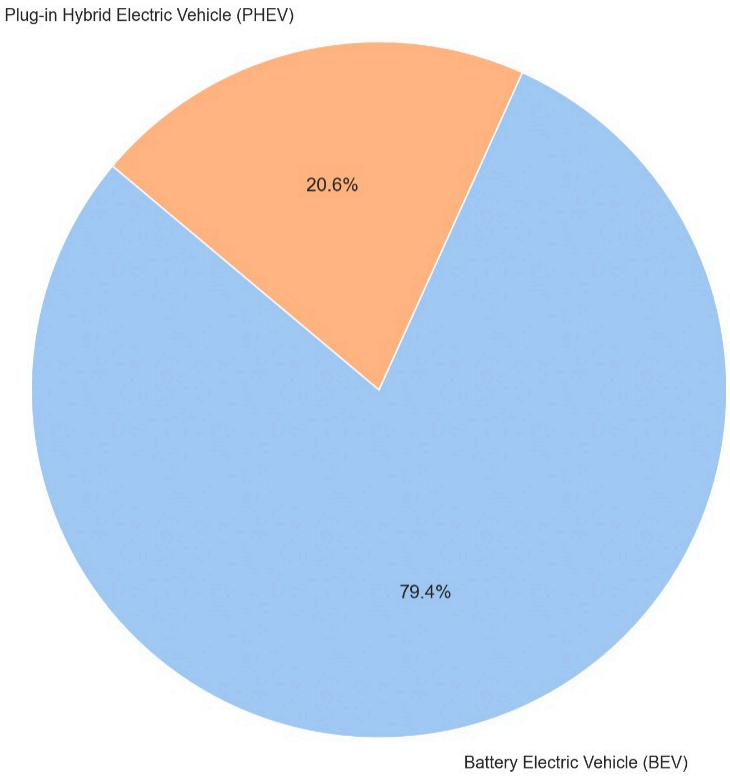
The dataset captures the number of EVs registered across various states, alongside associated infrastructure metrics like charging stations. By aggregating the EV counts state-wise, we can simulate a comparative trend that reveals regions leading in EV adoption and those that may require more support or infrastructure development.

4.1.3 Specific Requirements, Functions, and Formulas

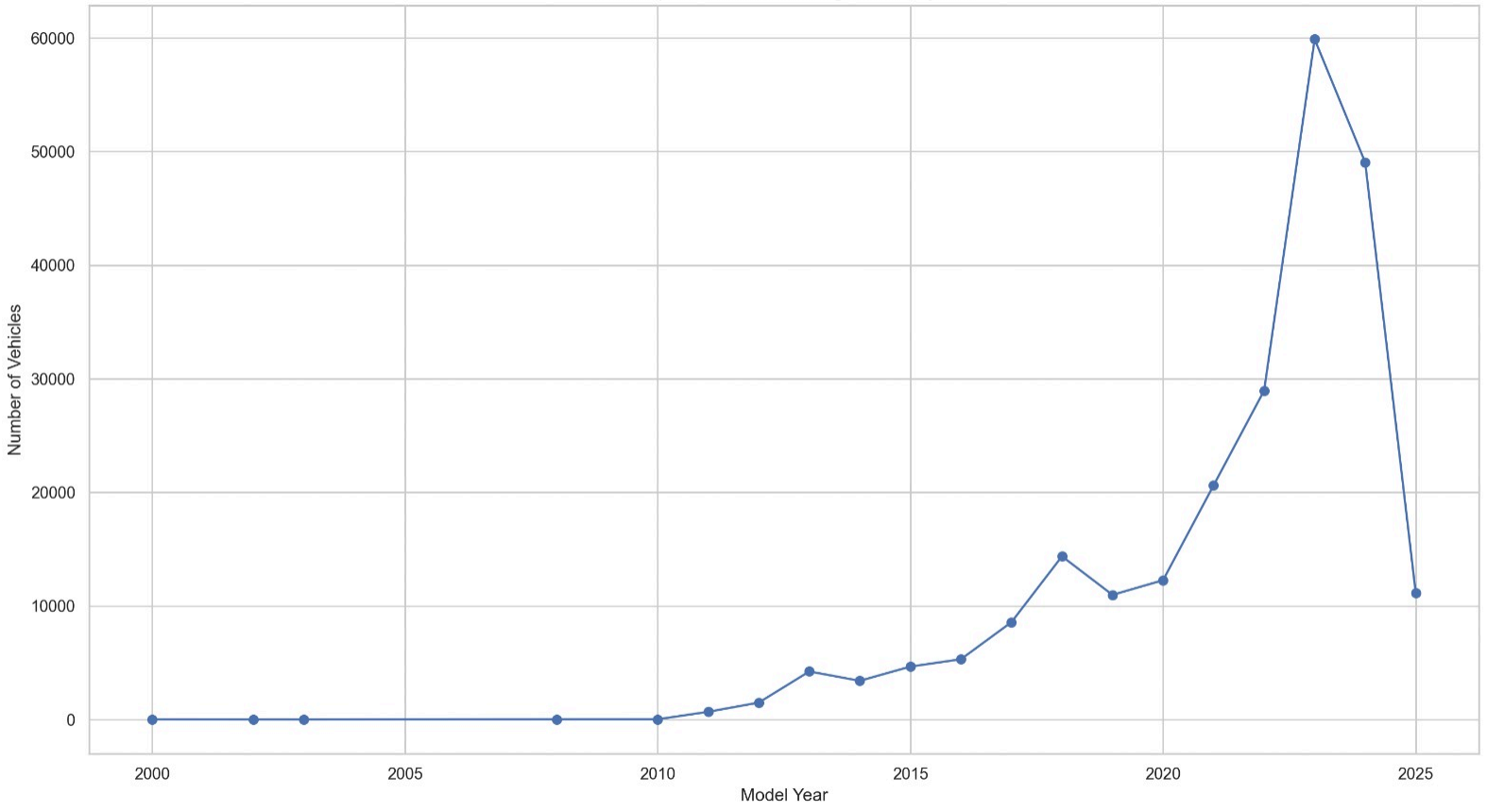
- **Libraries Used**: pandas, matplotlib.pyplot, numpy
- **Functions**: .groupby(), .sum(), plt.plot()
- **Formula Logic**: state_ev_population = df.groupby('State')['EV_Count'].sum()

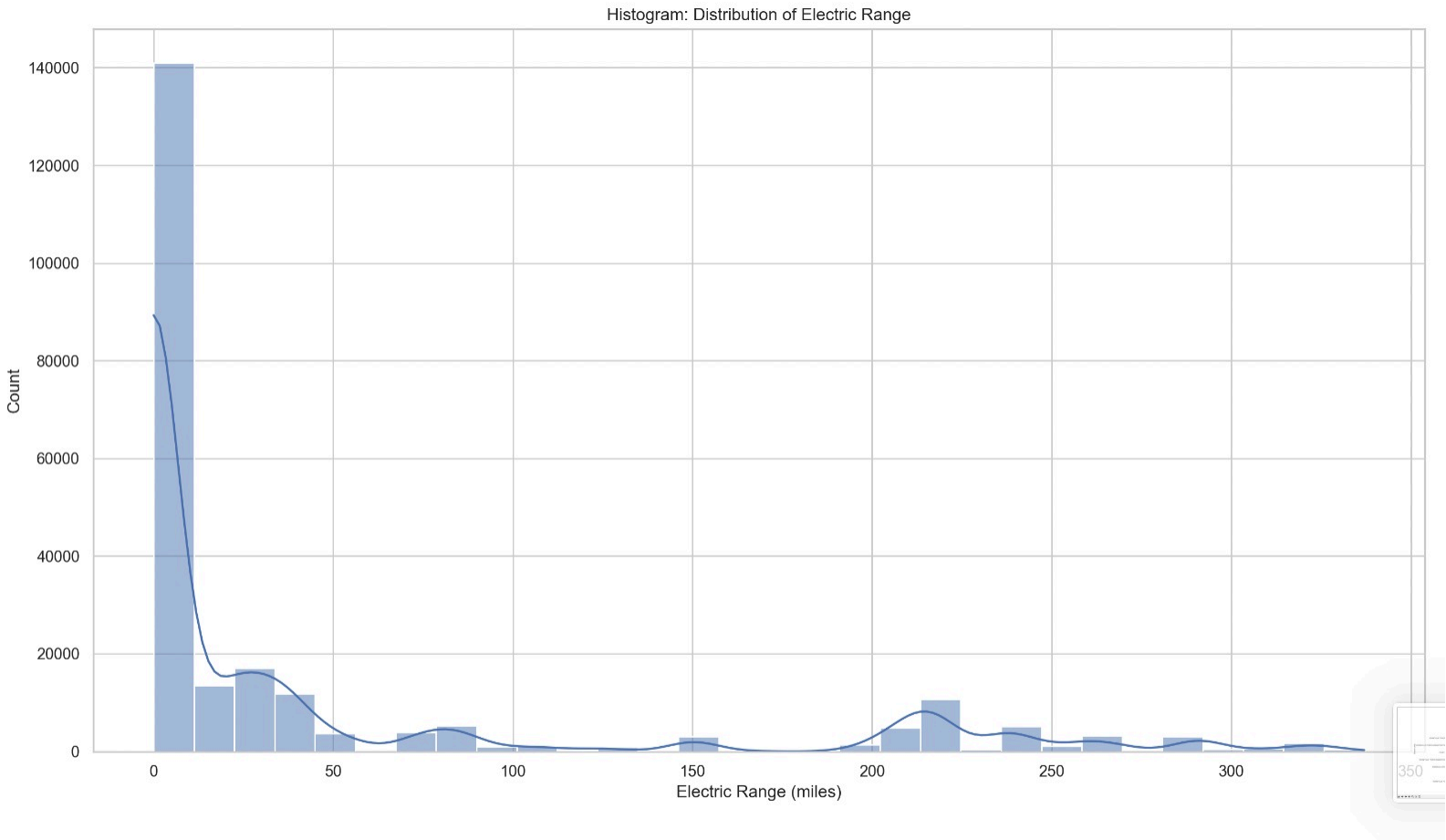
4.1.5 Visualization

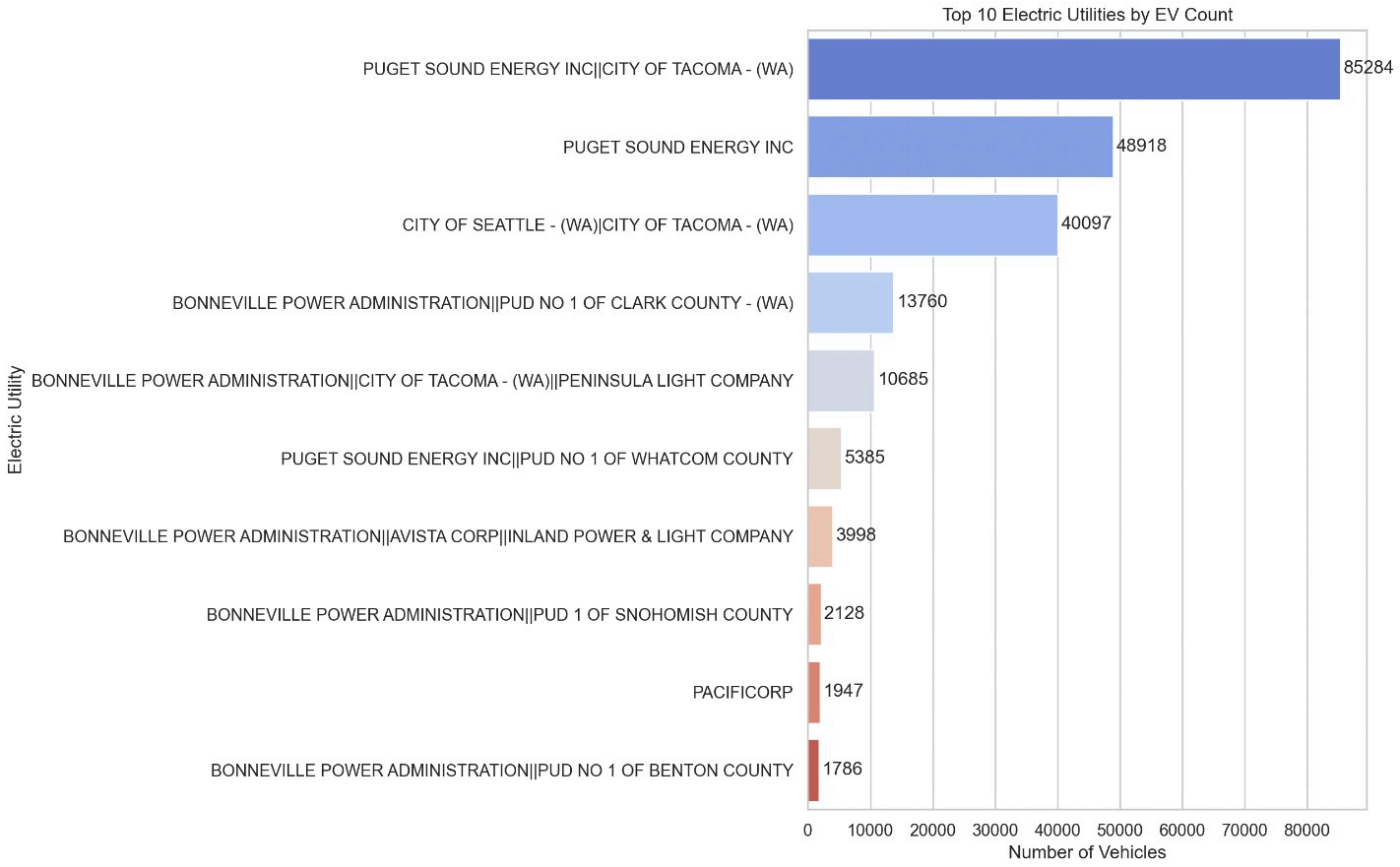
Pie Chart: Distribution of Electric Vehicle Types

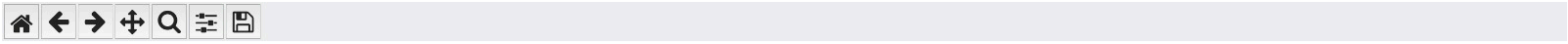


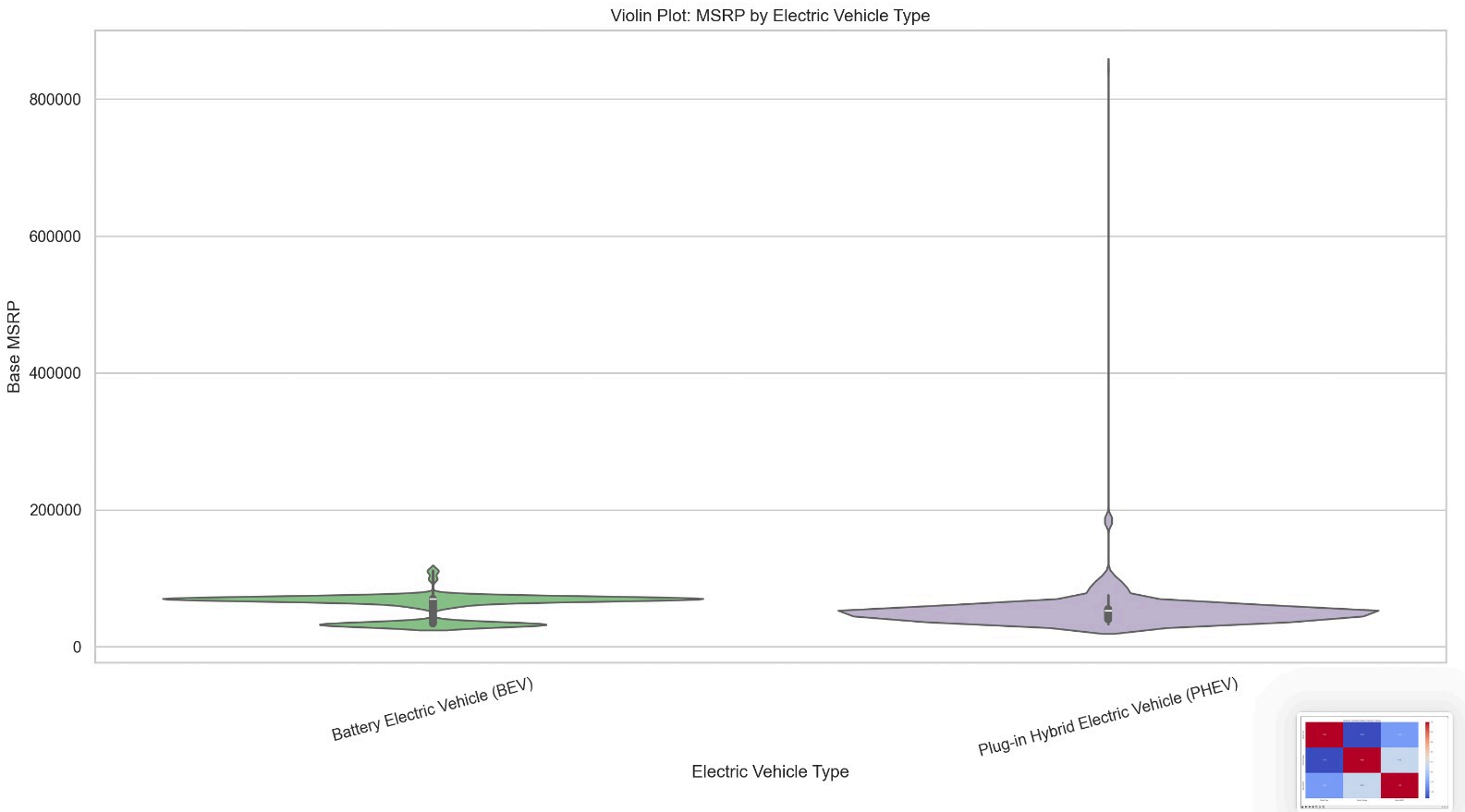
Line Graph: Electric Vehicle Registrations by Year

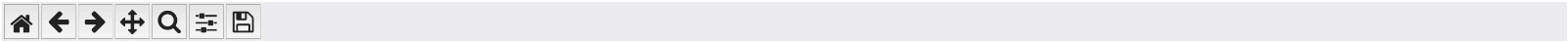


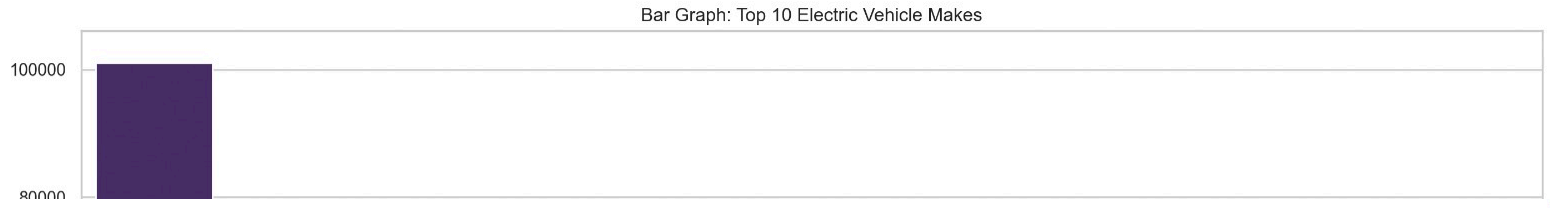

Histogram: Distribution of Electric Range
Count
140000
120000
100000
80000
60000
40000
20000
0
0
50
100
150
200
250
300
Electric Range (miles)

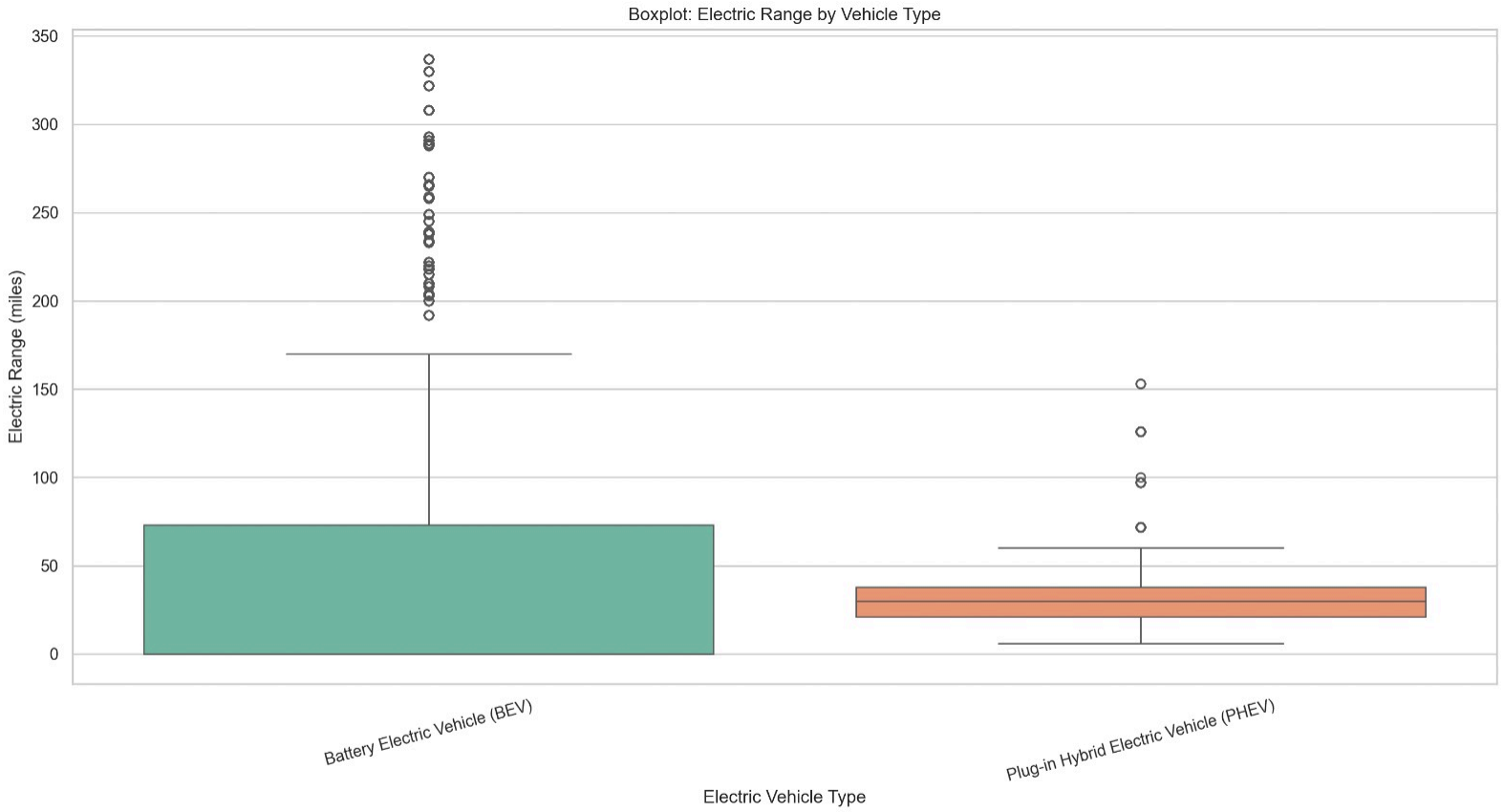

Top 10 Electric Utilities by EV Count
PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)
PUGET SOUND ENERGY INC
CITY OF SEATTLE - (WA)||CITY OF TACOMA - (WA)
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLARK COUNTY - (WA)
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PENINSULA LIGHT COMPANY
PUGET SOUND ENERGY INC||PUD NO 1 OF WHATCOM COUNTY
BONNEVILLE POWER ADMINISTRATION||AVISTA CORP||INLAND POWER & LIGHT COMPANY
BONNEVILLE POWER ADMINISTRATION||PUD 1 OF SNOHOMISH COUNTY
PACIFICORP
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF BENTON COUNTY
85284
48918
40097
13760
10685
5385
3998
2128
1947
1786
Electric Utility
0
10000
20000
30000
40000
50000
60000
70000
80000
Number of Vehicles

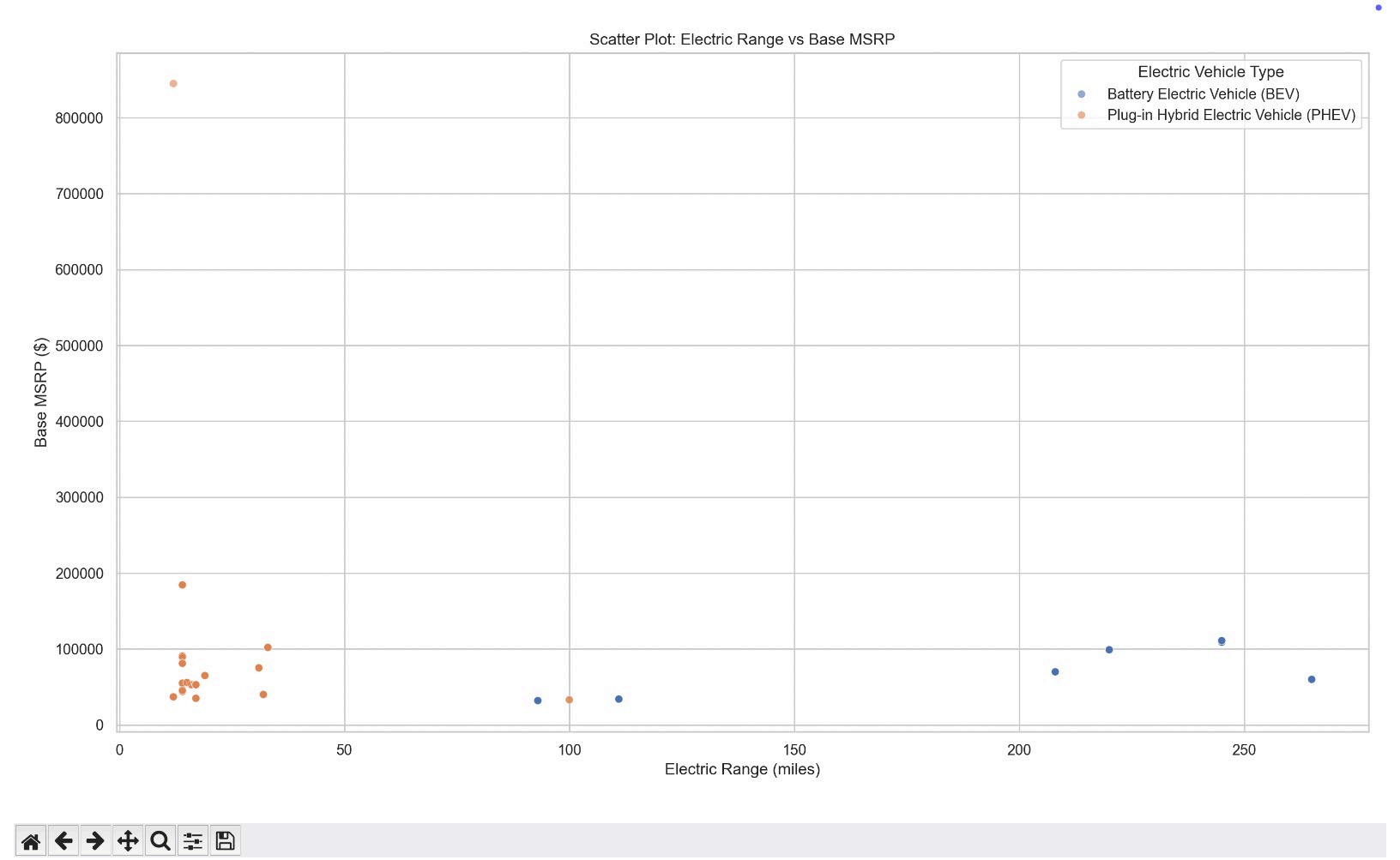


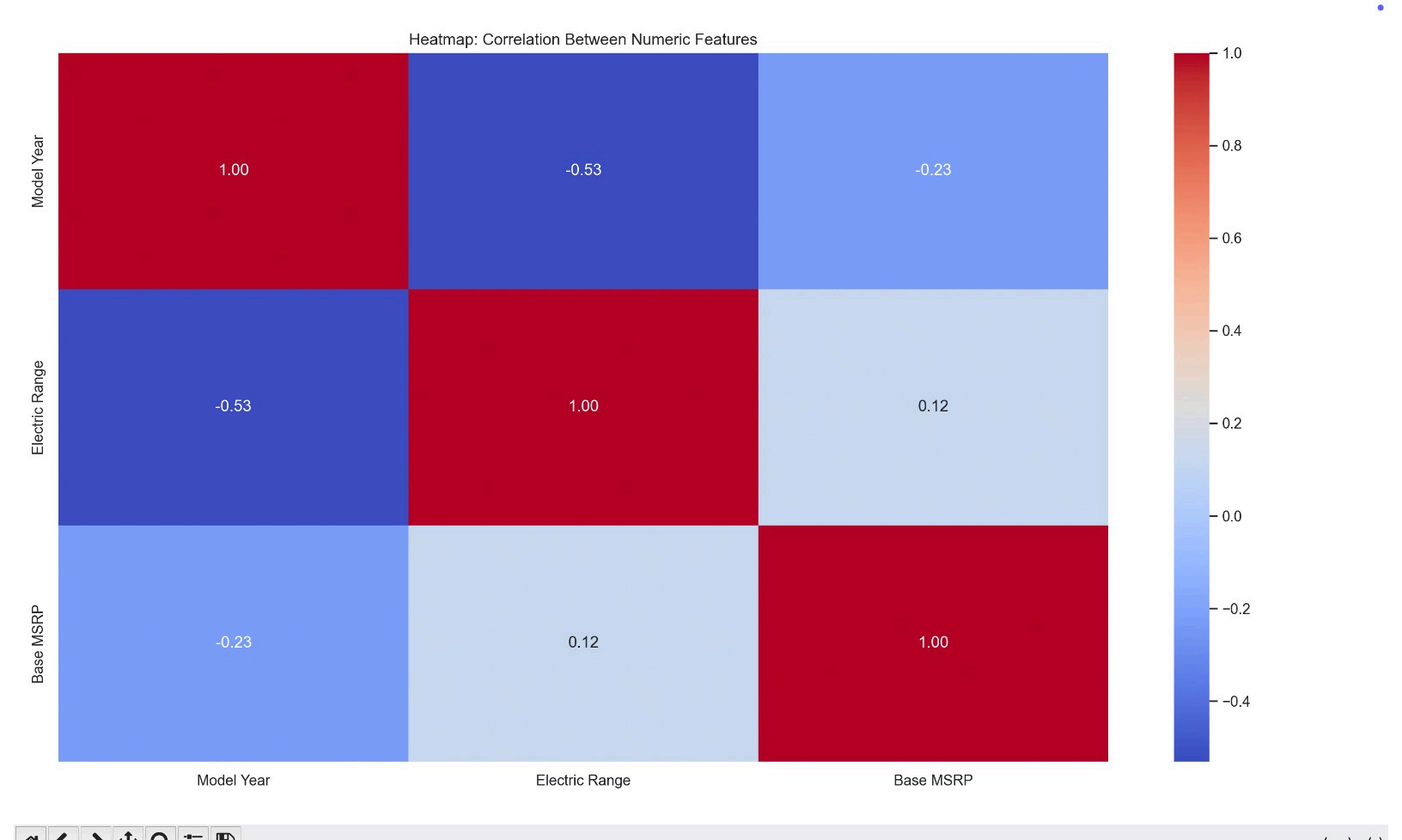

Violin Plot: MSRP by Electric Vehicle Type
800000
600000
400000
200000
0
Base MSRP
Battery Electric Vehicle (BEV)
Plug-in Hybrid Electric Vehicle (PHEV)
Electric Vehicle Type




Bar Graph: Top 10 Electric Vehicle Makes
100000
80000


Boxplot: Electric Range by Vehicle Type
Electric Range (miles)
350
300
250
200
150
100
50
0
Battery Electric Vehicle (BEV)
Plug-in Hybrid Electric Vehicle (PHEV)
Electric Vehicle Type


Scatter Plot: Electric Range vs Base MSRP
Electric Vehicle Type
Battery Electric Vehicle (BEV)
Plug-in Hybrid Electric Vehicle (PHEV)
Base MSRP ($)
800000
700000
600000
500000
400000
300000
200000
100000
0
0
50
100
150
200
250
Electric Range (miles)


Heatmap: Correlation Between Numeric Features
Model Year
Electric Range
Base MSRP
1.00
-0.53
-0.23
-0.53
1.00
0.12
-0.23
0.12
1.00
Model Year
Electric Range
Base MSRP
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

(x, y) = (,)

FUTURE SCOPE

The current analysis offers valuable insights into the distribution and adoption of electric vehicles (EVs) across various states and districts in India. However, the scope of this project can be significantly expanded in the following ways:

1. Temporal Comparisons
The dataset analyzed focuses on a specific year. Future work can incorporate multi-year EV registration data to examine growth trends, adoption rates, and policy impacts over time.

2. Integration with Demographic and Economic Factors
By combining EV data with population density, income levels, urban-rural ratios, and fuel prices, deeper insights can be drawn about factors influencing EV adoption in different regions.

3. Geospatial Analysis and Mapping
Utilizing Geographic Information System (GIS) tools can help create interactive maps showing EV density, charging infrastructure distribution, and infrastructure gaps, supporting more informed planning.

4. Predictive Modelling
Machine learning models can be developed to forecast EV adoption rates based on factors like state-wise subsidies, infrastructure expansion, and consumer behavior trends. This can aid policymakers in proactive planning.

5. Policy Planning and Infrastructure Allocation
Insights from this analysis can support government and private stakeholders in identifying underserved areas, guiding investments in charging stations, battery swap hubs, and service facilities.

6. Interactive Dashboards
Building real-time dashboards using tools such as Power BI, Tableau, or Dash can make the insights accessible to decision-makers and the public, promoting awareness and engagement in EV adoption.

7. Data Enrichment from External Sources
Integrating additional datasets such as road quality, traffic density, electricity grid capacity, and air quality indices can provide a more comprehensive view of the challenges and opportunities in expanding EV usage.

In conclusion, this project lays the groundwork for further research, strategic planning, and technology-driven innovation to accelerate India’s transition to clean and sustainable transportation.

REFERENCES

[1] Available: https://catalog.data.gov/dataset/electric-vehicle-population-data[Accessed: Apr. 10,2025].

[2] Wes McKinney, *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*, 2nd ed. O'Reilly Media, 2017.

[3] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, May–Jun. 2007.

[4] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proc. 9th Python in Science Conf. (SciPy 2010)*, Austin, TX, USA, 2010, pp. 51–56.

[5] NumPy Developers, *NumPy: The fundamental package for scientific computing with Python*, [online]. Available: https://numpy.org [Accessed: Apr. 10, 2025].

[6] Pandas Development Team, *Pandas Documentation*, [online]. Available: https://pandas.pydata.org [Accessed: Apr. 10, 2025].