

Greenhouse Gas Emissions Analysis 2024

Name: Papabaigari Mahammad Yaseen

Reg. No: 12317975

Programme: K23GN

Roll No: 12

Course Code: INT375

Guide: Mrs. Aashima

Lovely Professional University

Introduction

Explores GHG emissions for Sept 2024 quarter.

Includes industry and household data.

Uses Python for analysis and visualization.

Focuses on Environmental Data Understanding (EDU).

EDU aids in sustainable decisions.

Highlights patterns, outliers, and trends.

Objective: extract insights from emissions data.

Supports decisions in climate strategies.

Source of Dataset

Dataset: Greenhouse Gas Emissions - Sept 2024.

Source: Stats NZ (New Zealand).

Format: CSV.

Location: Local system path.

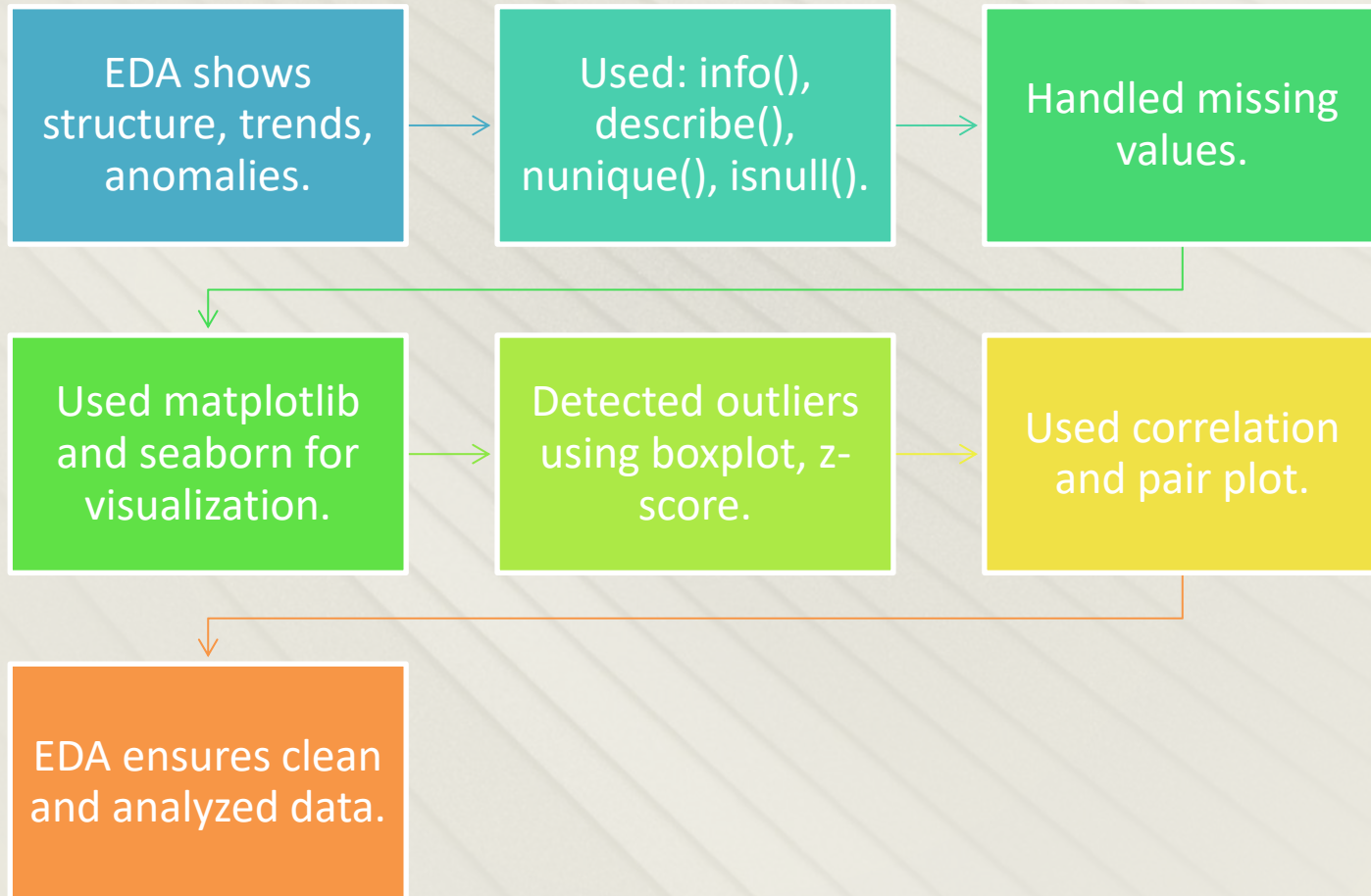
Includes household and industry data.

Measured in kilotonnes.

Has both time and sector data.

Core input for this analysis.

EDA Process Overview



Python Tools Used

Pandas: Data manipulation.

NumPy: Numerical operations.

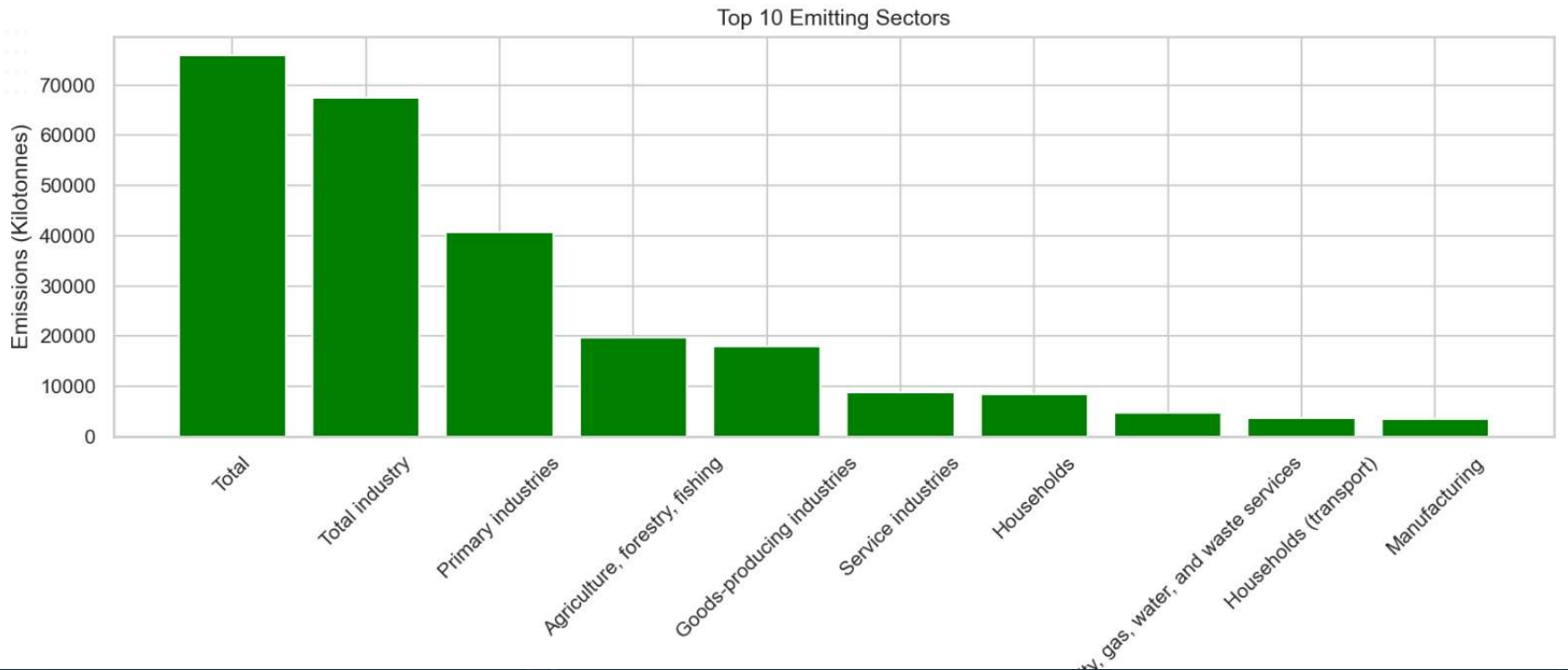
Matplotlib: Basic visualizations.

Seaborn: Advanced visualizations.

SciPy: Z-score for outlier detection.

Libraries ensure effective EDA.

Create clean, visual, interpretable data.



Top Emitting Sectors (Bar Plot)

- Shows sectors with highest emissions.
- Uses latest period data.
- Grouped and summed emissions.
- Top 10 sectors displayed in bar chart.
- Functions: `groupby()`, `sum()`, `sort_values()`.
- Highlights major contributors.
- Useful for sector-specific decisions.

Household Emissions Trend (Line Graph)

Tracks household emissions over time.

Grouped by time period.

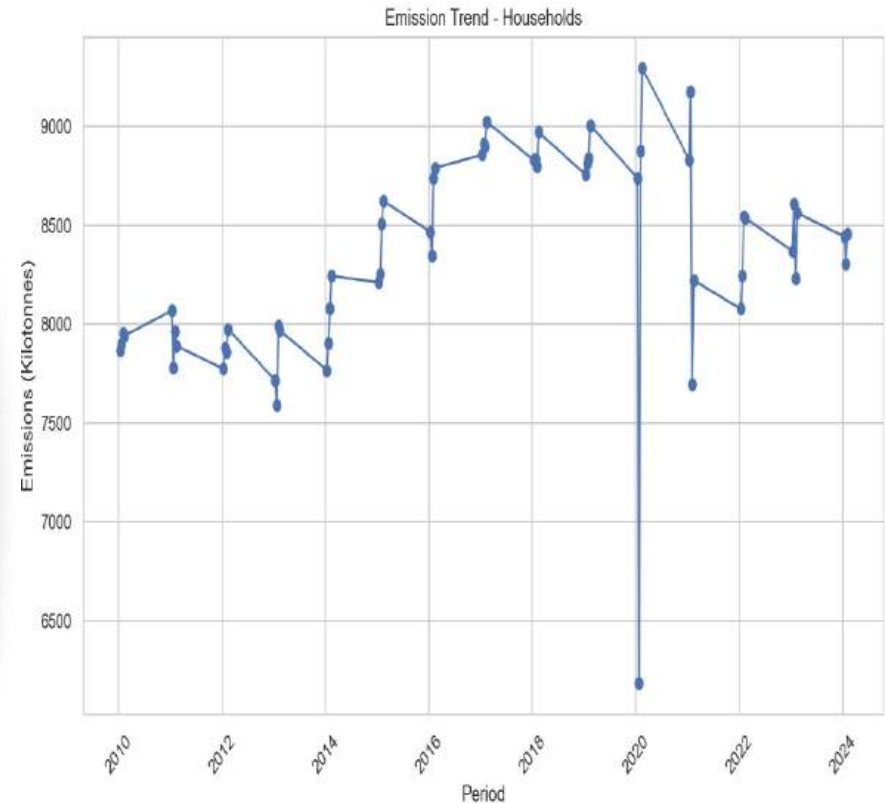
Visualized using line graph.

Shows trend: increasing/decreasing.

Functions: `groupby()`, `plot()`.

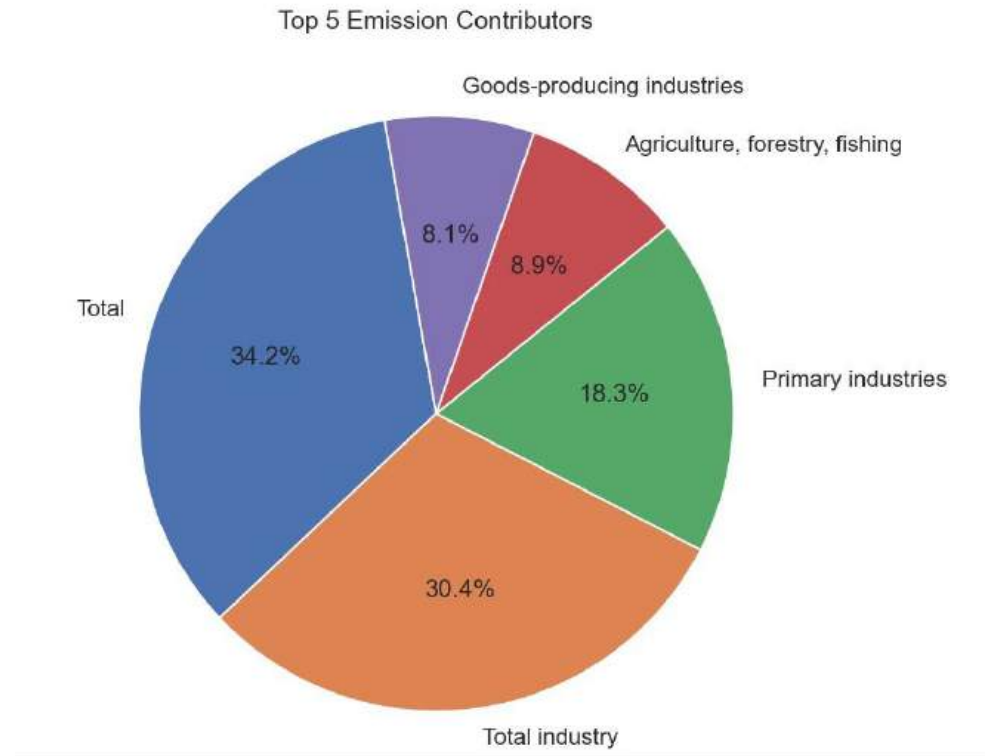
Indicates public impact.

Helps shape awareness strategies.



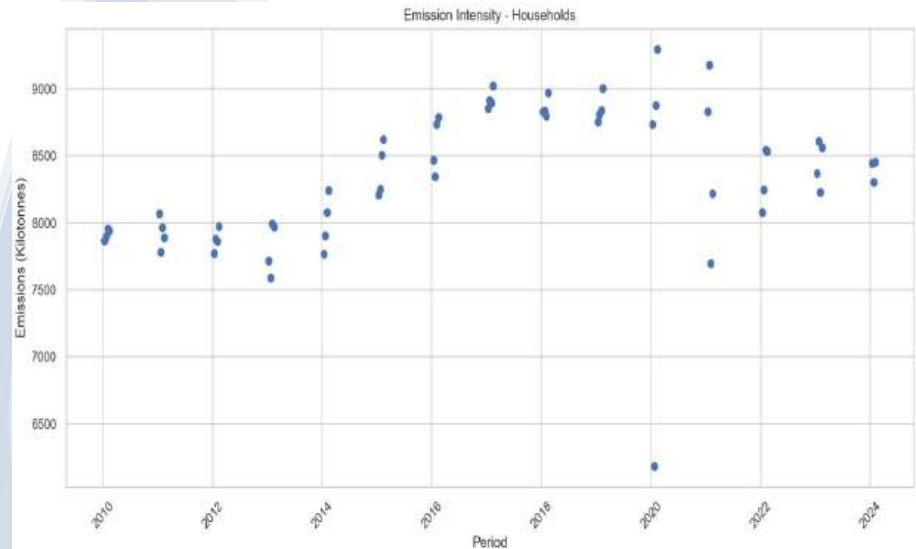
Top 5 Contributors (Pie Chart)

- Pie chart for top 5 emission sectors.
- Grouped and sorted data.
- Functions: `groupby()`, `sum()`, `pie()`.
- Displays contribution shares.
- Identifies major polluters.
- Supports sector-focused policies.



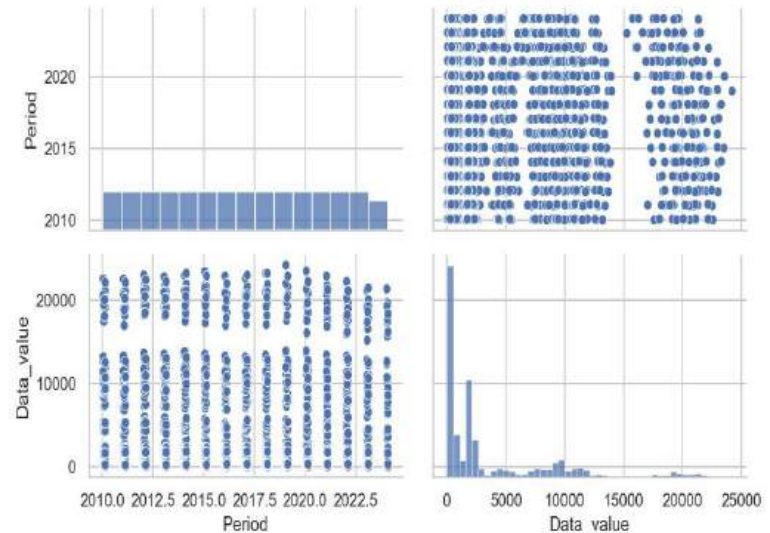
Emission Intensity (Scatter Plot)

- Scatter plot of household emissions.
- Shows each period's values.
- Spikes and drops visualized.
- Grouped data by period.
- Functions: `groupby()`, `scatter()`.
- Highlights household impact trends.



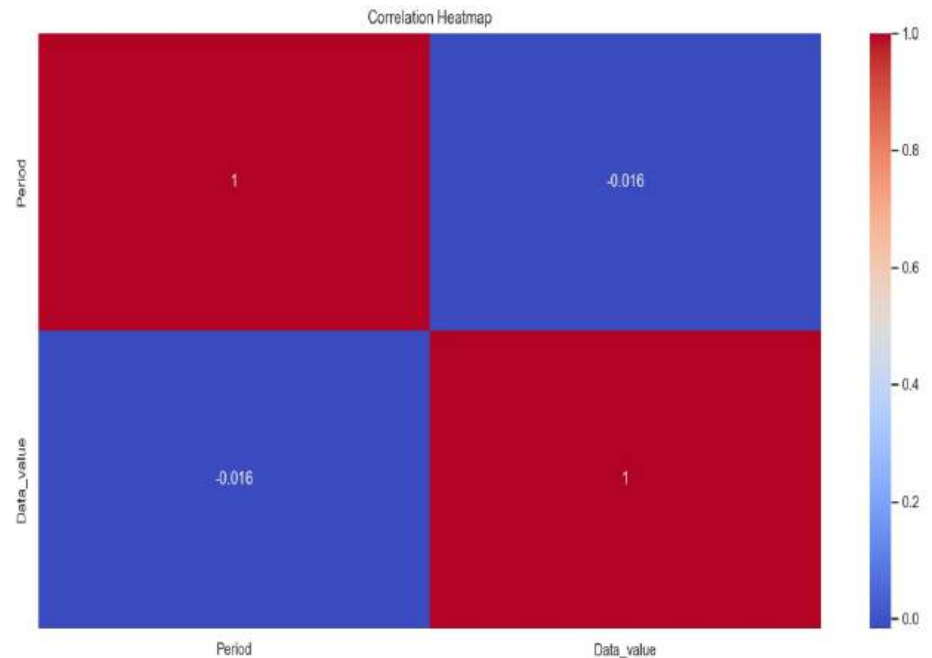
Pair Plot of Variables

- Shows pairwise variable relationships.
- Used seaborn's pairplot().
- Detects clusters, trends.
- Supports multivariate analysis.
- Helps interpret data distribution.
- Initial step for modeling.



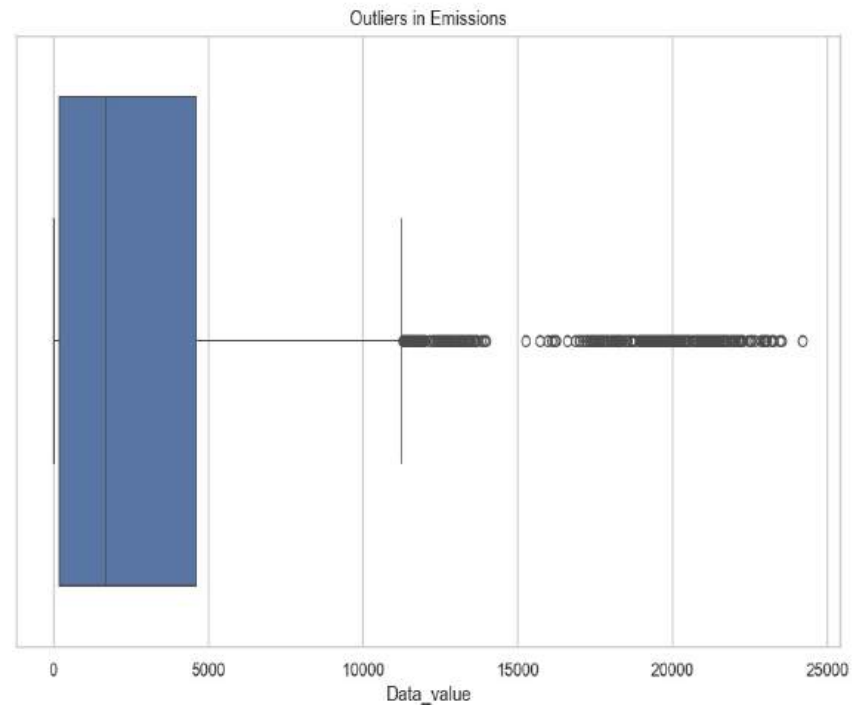
Correlation Heatmap

- Visualizes correlation between variables.
- Used `corr()` and `heatmap()`.
- Shows strength of linear relationships.
- Color-coded for easy reading.
- Supports feature selection.
- Reduces multicollinearity.



Boxplot for Outliers

- Detects extreme values visually.
- Used seaborn boxplot.
- Emissions column analyzed.
- Shows median, spread, and outliers.
- Supports data cleaning.
- Quick outlier identification.



Z-Score Based Outlier Detection

Quantitative
outlier method.

Uses SciPy's
`zscore()`.

$Z > 3$ indicates
outliers.

Complementary
to boxplot.

Finds anomalies
statistically.

Ensures data
reliability.

Conclusion



Revealed top sectors for emissions.



Analyzed trends and intensity.



Detected key outliers.



Used multiple visualization tools.



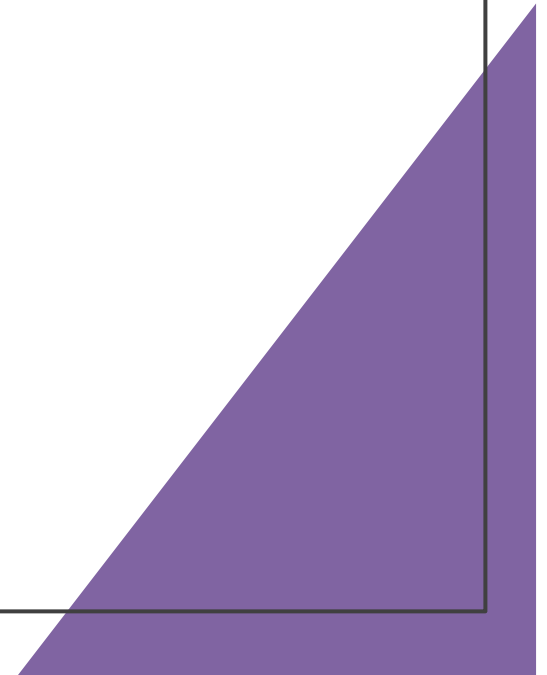
Strengthened data understanding.



Supports eco-policy decisions.

Future Scope

- Predict emissions using ML.
- Integrate weather/population data.
- Automate real-time dashboards.
- Apply clustering on industries.
- Expand to global datasets.
- Enhance public policy transparency.



The background features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green, creating a modern and dynamic look. The shapes are concentrated on the right side of the slide, with some extending towards the left.

Thank You