CO2 Emission Prediction From Vehicles

Team Members

Dhiraj Mahato

Rohit Dey

Anudipta Das

Ishita Karmakar

Rishov Ghosh

Sayed Rahaman

Overview

- Introduction
- Problem Identification
- Tools used for Analysis
- Data Collection
- Data Cleaning
- Data Analysis
- Data Modeling
- Hypothesis Development
- Solution Design
- Visualization (Power BI)
- Conclusion
- References

Introduction

- This project is a data-driven initiative with which we aim to completely analyze the complex interplay between various factors (e.g., engine size, cylinder count, fuel consumption) and CO2 emissions from different vehicles.
- Our goal is to develop a machine learning model that generates precise CO2 emission predictions based on vehicle attributes.
- In order to comprehend the effects of regulatory changes and technology improvements on environmental sustainability, the primary objective of this study is to analyze trends in vehicle CO2 emissions.

Problem Identification

Problem Statement

- The CO_2 emissions from vehicles have a major impact on both climate change and greenhouse gas emissions worldwide.
- In order to anticipate emissions in future vehicle designs, it is imperative to comprehend the components controlling vehicle CO_2 emissions.
- In order to make well-informed decisions, research examines CO_2 emissions across different vehicle types.

Significance of the Problem

- Climate change mitigation: Reducing vehicle emissions is crucial for mitigating global CO_2 emissions.
- Air quality and public health: Lowering vehicle emissions can improve air quality and health outcomes.
- Policy development: Data-driven insights inform fuel efficiency policies and incentives for low-emission vehicles.
- Consumer awareness: Clear vehicle emissions information aids environmentally conscious purchasing decisions.
- Automotive industry impact: Understanding emission trends drives innovation towards fuel-efficient and low-emission vehicle technologies.

Relevant SDGs

SDG 3: Good Health and Well-Being

- Reducing vehicle emissions improves air quality, directly impacting public health.

SDG 7: Affordable and Clean Energy

- The project relates to improving energy efficiency in transportation.

SDG 9: Industry, Innovation and Infrastructure

- Insights from the analysis can drive innovation in automotive technology and transportation infrastructure.

SDG 11: Sustainable Cities and Communities

- Reducing vehicle emissions is key to creating more sustainable urban environments.

SDG 13: Climate Action

- Addressing vehicle CO_2 emissions is directly relevant to climate change mitigation efforts.

Tools used for Analysis

» 01
Python

For data cleaning, analysis, and visualization, using libraries such as Pandas, NumPy, Matplotlib, Seaborn and Scikit-learn.

» 02
Jupyter Notebook

For documenting the analysis process and visualizations.

» 03
Power BI

For developing interactive visualizations and dashboards.

Libraries used in Python

1. Pandas
2. NumPy
3. Matplotlib
4. Seaborn
5. Scikit-learn

Data Collection

Name of Dataset: CO2 Emission by Vehicles

Source: Kaggle

Link: Click Here

Snapshot of Dataset:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Make	Model	Vehicle	Engine	Cylinders	Transm	Fuel Typ	Fuel	Fuel (	Fuel	Fuel (	CO2 Emi
2	ACURA	ILX	COMP/	2	4	AS5	Z	9.9	6.7	8.5	33	196
3	ACURA	ILX	COMP/	2.4	4	M6	Z	11	7.7	9.6	29	221
4	ACURA	ILX HYE	COMP/	1.5	4	AV7	Z	6	5.8	5.9	48	136
5	ACURA	MDX 4\	SUV -	3.5	6	AS6	Z	13	9.1	11	25	255
6	ACURA	RDX A\	SUV -	3.5	6	AS6	Z	12	8.7	11	27	244

Data Cleaning

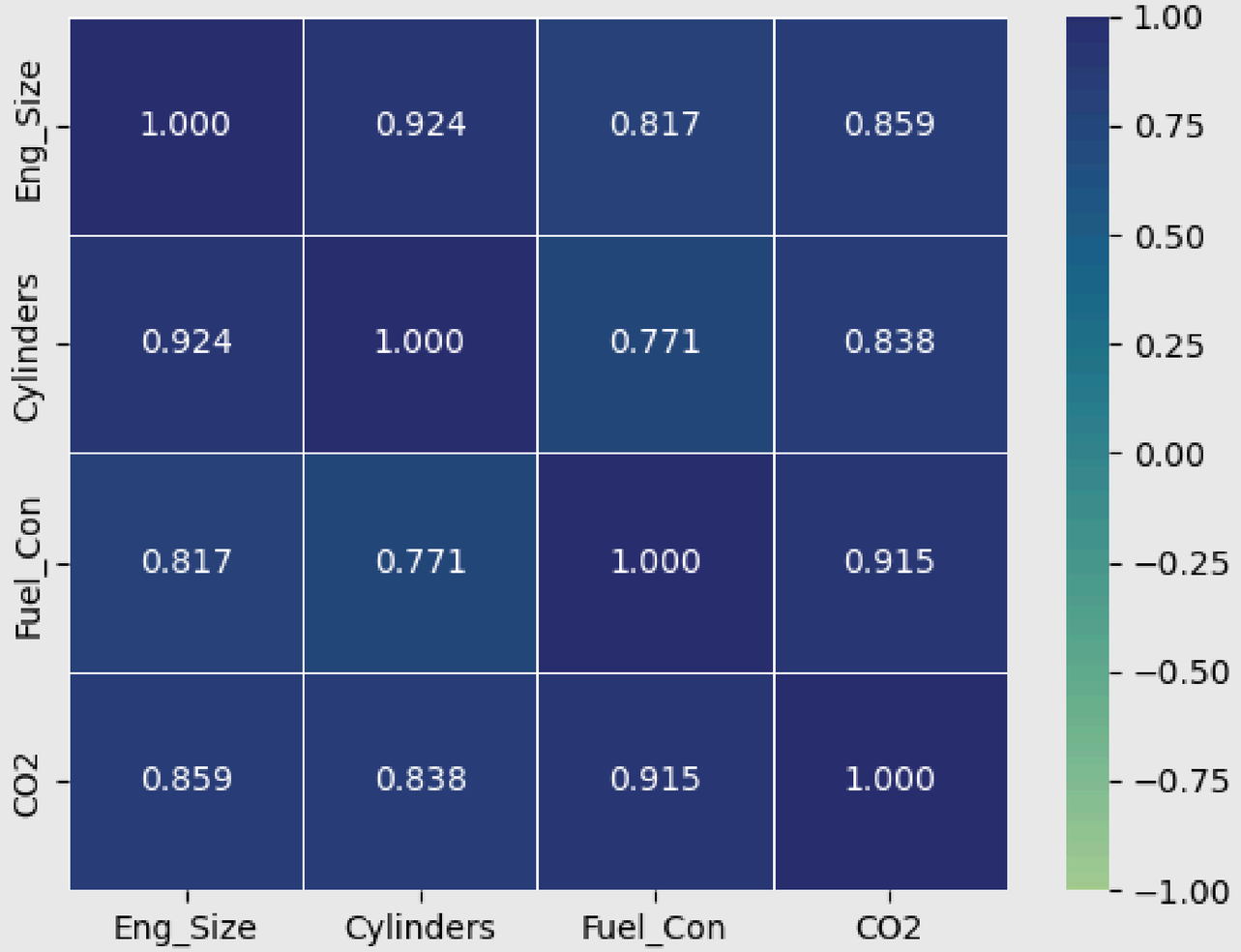
- Dropping Unnecessary Columns
- Formatting Model Column Values to Upper Case
- Dropping Duplicates

Snapshot of Cleaned Dataset:

	A	B	C	D	E	F	G	H	I
1	Make	Model	ehicle Clas	gine Size(	Cylinders	ransmissid	Fuel Type	ption Com	missions(
2	ACURA	ILX	COMPACT	2	4	AS5	Z	8.5	196
3	ACURA	ILX	COMPACT	2.4	4	M6	Z	9.6	221
4	ACURA	ILX HYBRI	COMPACT	1.5	4	AV7	Z	5.9	136
5	ACURA	MDX 4WD	SUV - SM.	3.5	6	AS6	Z	11.1	255
6	ACURA	RDX AWD	SUV - SM.	3.5	6	AS6	Z	10.6	244
7	ACURA	RLX	MID-SIZE	3.5	6	AS6	Z	10	230

Data Analysis

- **Correlation heatmap of the numerical features**
- **Parameters needed for CO2 emission prediction in decreasing order of importance are:**
  - **Fuel Consumption**
  - **Engine Size**
  - **Cylinders**



Random Forest Model

Attributes Taken:

- Fuel Consumption
- Engine Size
- Cylinder Count

Model Accuracy:

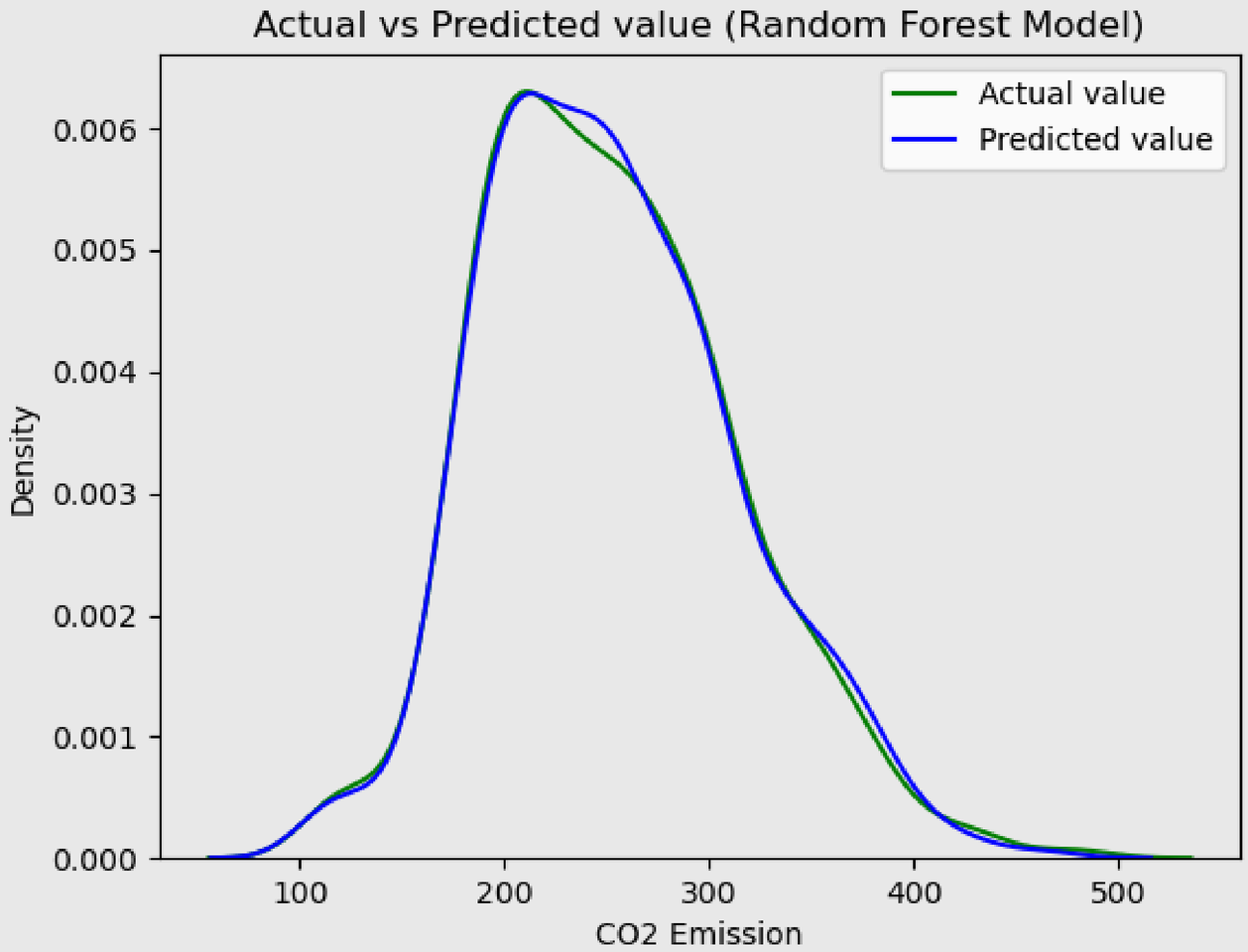
- 96.59%

Training Sample Size:

- 2173

Testing Sample Size:

- 932



Hypothesis Development

Adoption of advanced fuel-efficient technologies and increased EV use:

- **Substantially decrease CO_2 emissions from transportation over the years.**

Detailed analysis will:

- **Reveal specific temporal and spatial patterns in CO_2 emissions.**

These patterns can be mitigated through:

- **Focused policy measures.**
- **Targeted infrastructure development**

Solution Design

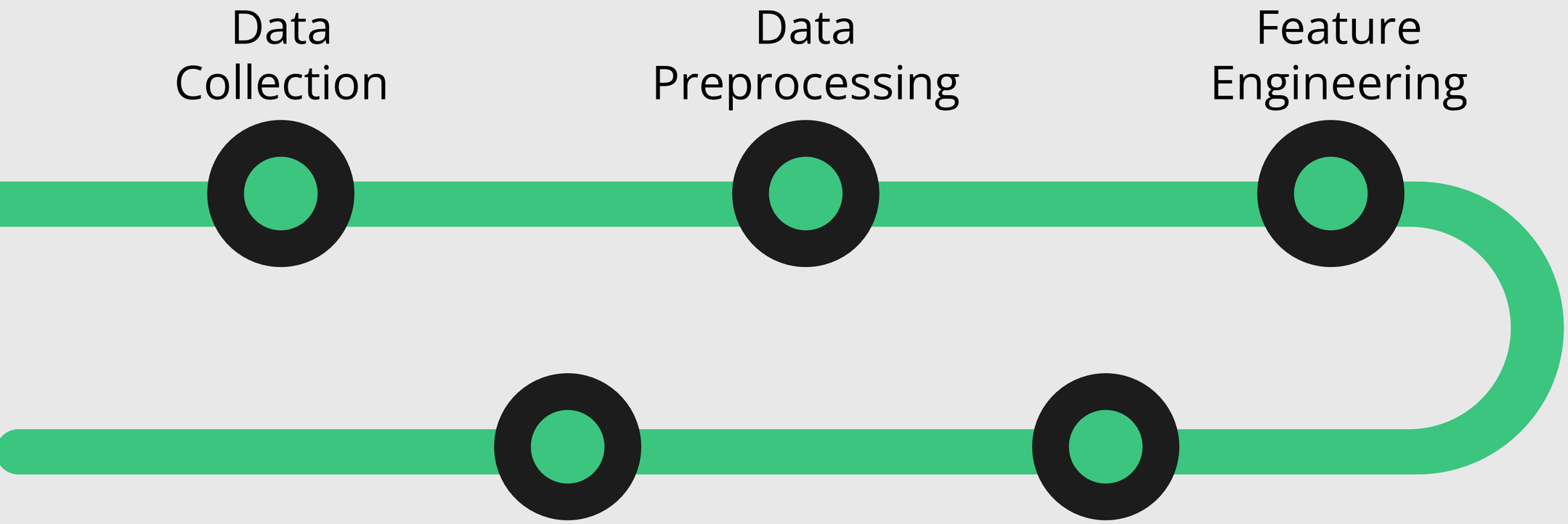
- Proposed Solution
- Implementation Plan
- Alignment with SDG’s

01 Proposed Solution

Objective: Develop a machine learning model to predict CO2 emissions based on vehicle mpg, cylinder count, and engine size.

Impact: This solution helps manufacturers design more fuel-efficient vehicles, aids policymakers in enforcing environmental regulations, and raises public awareness about emissions.

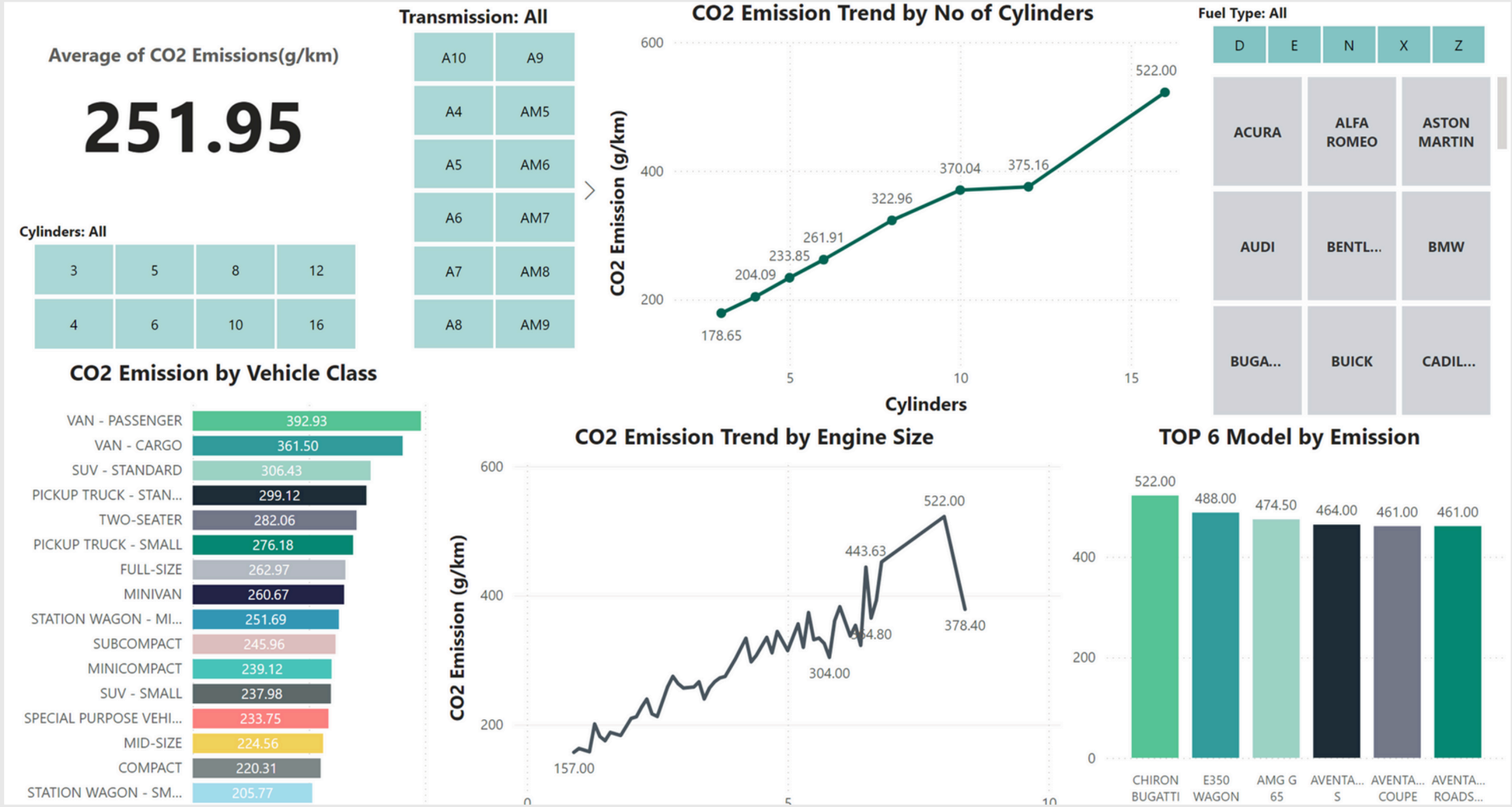
02 Implementation Plan



03 Alignment with SDGs

- **SDG 3:** Good Health and Well-Being
- **SDG 7:** Affordable and Clean Energy
- **SDG 9:** Industry, Innovation and Infrastructure
- **SDG 11:** Sustainable Cities and Communities
- **SDG 13:** Climate Action

Power BI Dashboard



[Dashboard link]

Conclusion

In conclusion, this project marks a crucial step toward a sustainable automotive industry through data-driven CO2 emission analysis and solutions. It promises to influence vehicle design, policy-making, and public awareness, driving significant reductions in automotive emissions.

In future, we want to expand our work to include more vehicle types, integrate real-time data, and explore AI-driven predictive models. Potential also exists for synergies with smart city initiatives to further optimize emissions reduction in the transport sector. We also integrate the model as a Website.

Project Link (Github)

References

Data source:

- **Kaggle**

Tools used:

- **Python**
- **Jupyter Notebook**
- **Power BI**

THANK YOU