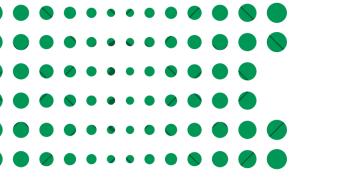


Regression Analysis of Real-World CO₂ Emissions



Exploring Factors Influencing Vehicle Emissions

By Shreya Mishra



Overview of the Dataset and Data Dictionary

Source: U.S. Environmental Protection Agency (EPA)

Dataset: EPA's Automotive Trends Report

Coverage:

Data spans from model year 1975 - 2009

Relevance:

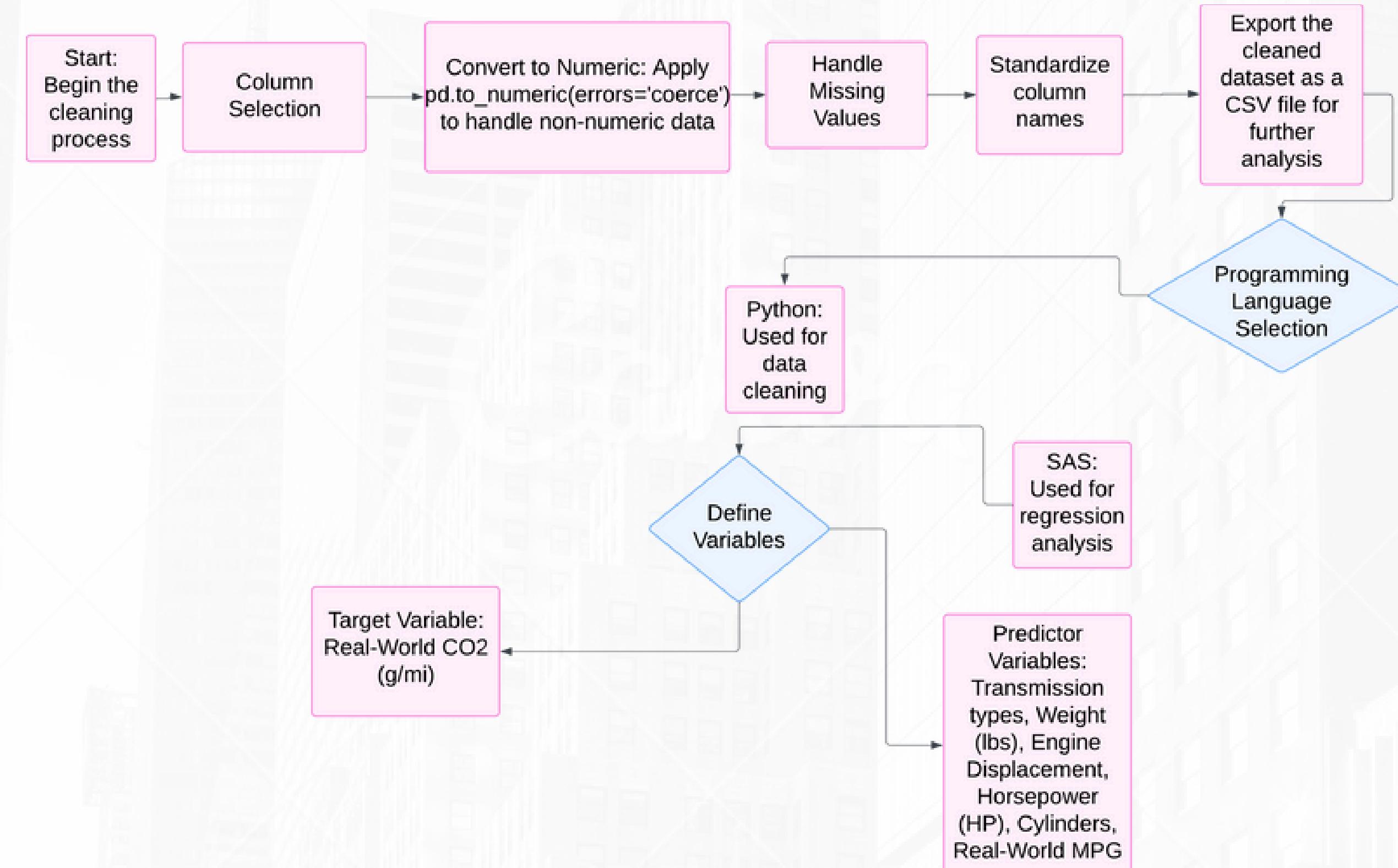
Ideal for studies on the impact of vehicles on climate change

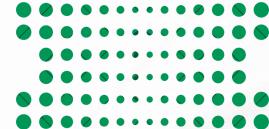
Dataset Suitability for Linear Regression:

Rich in quantitative variables relevant to vehicle performance and environmental impact.

Variable Name	Description	Type
Transmission_0_Automatic	Proportion or flag indicating if the vehicle has an automatic transmission.	Numeric (Continuous/Binary)
Transmission_0_Manual	Proportion or flag indicating if the vehicle has a manual transmission.	Numeric (Continuous/Binary)
Weight (lbs)	The total weight of the vehicle in pounds.	Numeric (Continuous)
Engine Displacement	The volume, in cubic centimeters, swept by all the pistons of the engine.	Numeric (Continuous)
Horsepower (HP)	The amount of power the engine produces.	Numeric (Continuous)
Cylinders	The number of cylinders in the vehicle's engine.	Numeric (Continuous/Discrete)
RealWorld MPG	The miles per gallon the vehicle achieves in real-world driving conditions.	Numeric (Continuous)
RealWorld CO2 (g/mi)	The amount of carbon dioxide the vehicle emits in grams per mile, measured under real-world driving conditions.	Numeric (Continuous)
Manufacturer	The company that manufactures the vehicle.	Categorical (Nominal)
Model Year	The year of the vehicle model.	Numeric (Discrete)
Regulatory Class	Classification of the vehicle according to regulatory guidelines.	Categorical (Nominal)
Vehicle Type	The type of vehicle, such as SUV, sedan, truck, etc.	Categorical (Nominal)

Data Cleaning and Analysis Framework





Model Results and Insights

Model Performance:

- R-Square: 0.9872 - Explains 98.7% of variation in CO2 emissions.

Significant Predictors:

- Transmission Type
- Weight & Engine Displacement
- Horsepower
- MPG : Miles Per Gallon

Statistical Significance:

- Model is highly significant ($p < 0.0001$), ensuring reliability of the results.
- Efficient fuel usage (MPG) is pivotal in lowering CO2 emissions.

Environmental Impact:

- Highlights how vehicle design modifications can mitigate environmental impacts.

Model: MODEL1
Dependent Variable: Real0World CO2 (g/mi)

Number of Observations Read	4162
Number of Observations Used	4162

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	132103877	18871982	45816.0	<.0001
Error	4154	1711067	411.90833		
Corrected Total	4161	133814944			

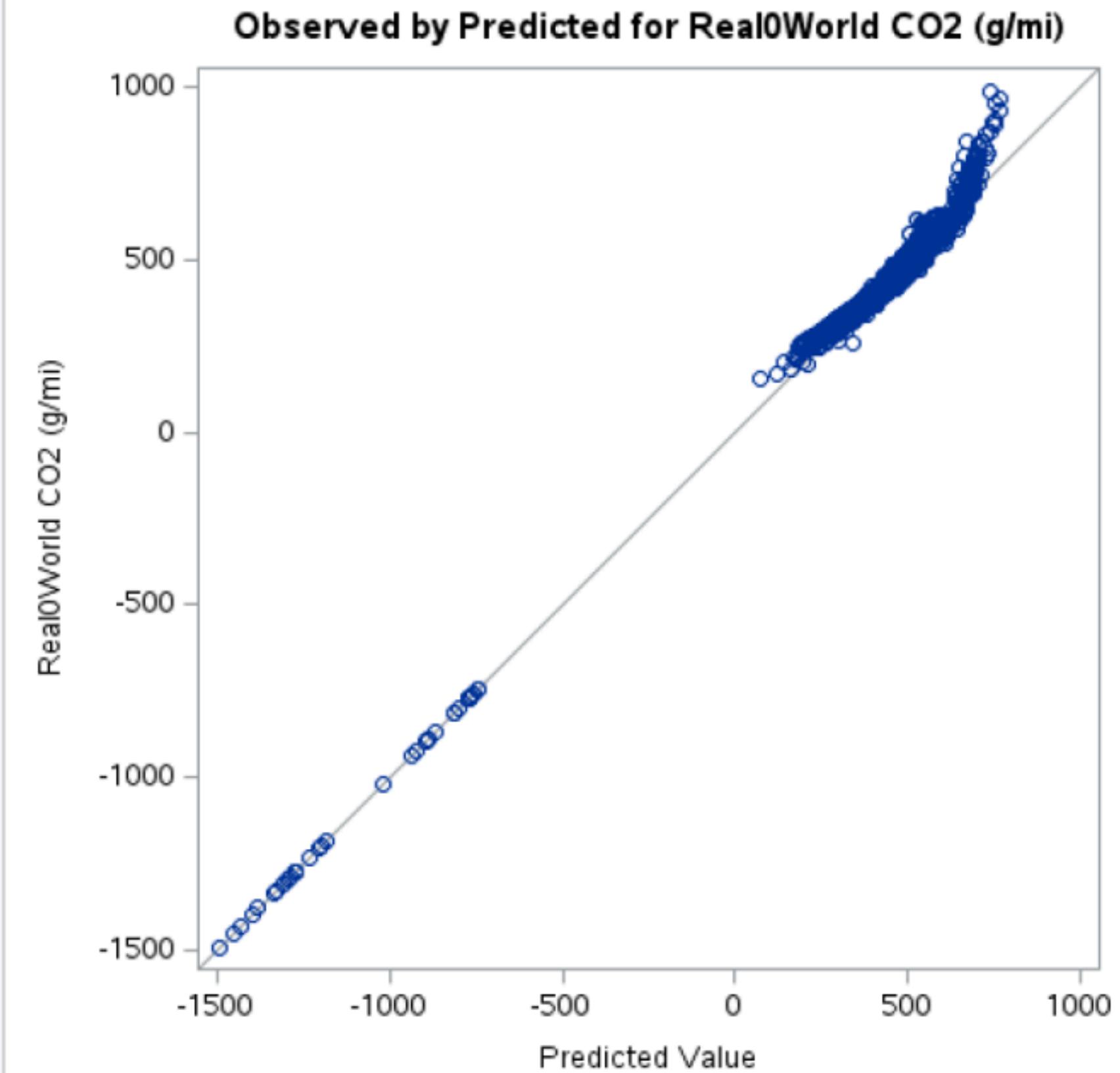
Root MSE	20.29552	R-Square	0.9872
Dependent Mean	413.72022	Adj R-Sq	0.9872
Coeff Var	4.90562		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	581.12198	6.52794	89.02	<.0001
Transmission 0 Automatic	1	81.09722	1.94597	41.67	<.0001
Transmission 0 Manual	1	59.19794	2.26210	26.17	<.0001
Weight (lbs)	1	0.02291	0.00109	21.02	<.0001
Engine Displacement	1	0.40748	0.01787	22.80	<.0001
Horsepower (HP)	1	-0.09421	0.01218	-7.73	<.0001
Cylinders	1	-1.61981	0.91933	-1.76	0.0782
Real0World MPG	1	-14.51743	0.11290	-128.59	<.0001



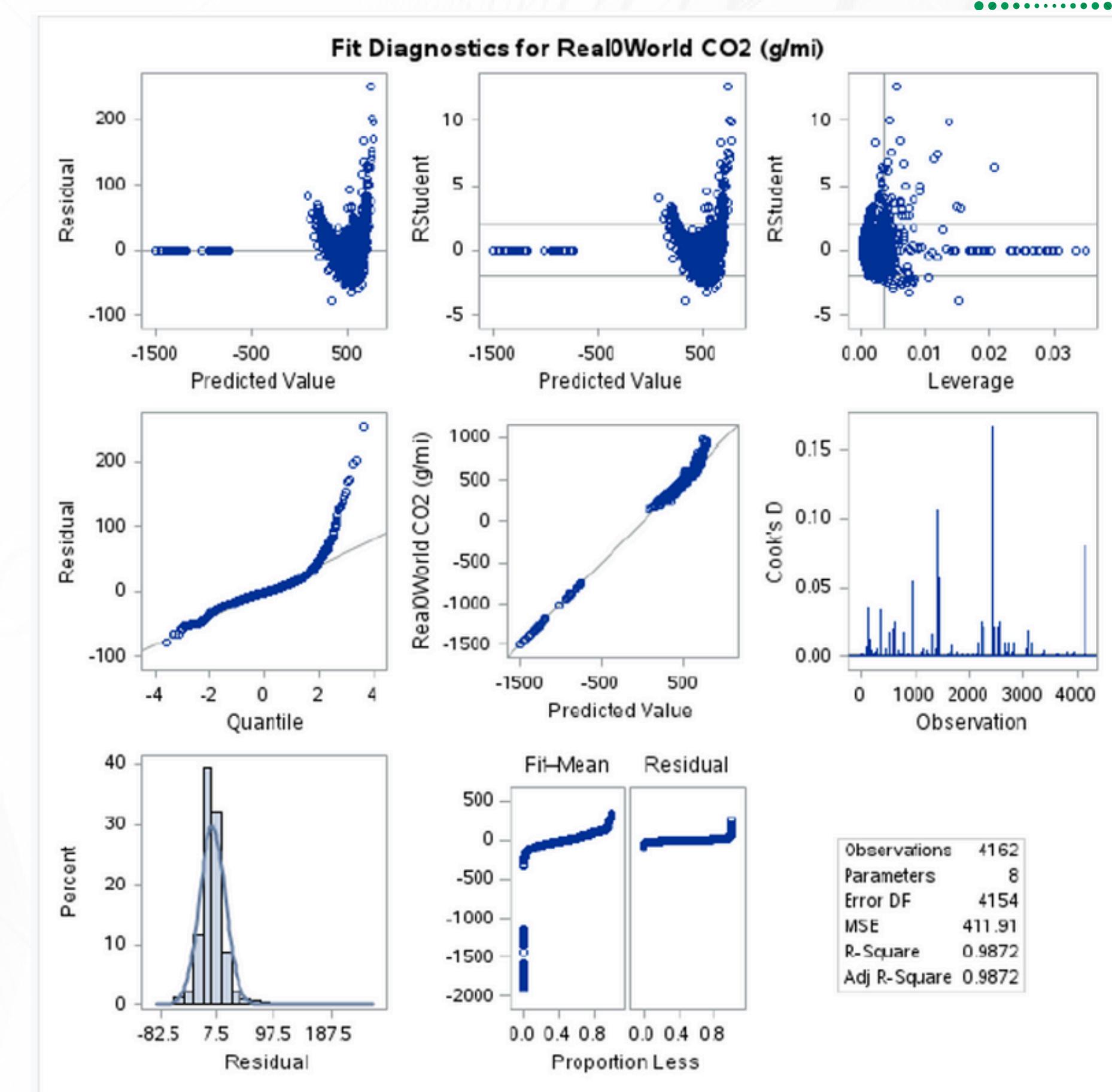
Key Insights from the Observed vs. Predicted Plot - CO2 Emission

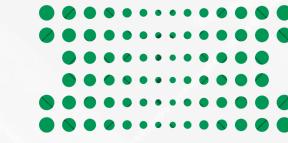
- Strong Observed-Predicted CO2 Correlation
- **High Accuracy:** Data Points Align with Diagonal
- **Low Residual Errors:** Minimal Deviation Observed
- R-Square 0.9872 Confirms **Strong Predictive Performance**
- **Limited Outliers** with Minimal Impact on Fit



Key Insights from Fit Diagnostics Plots

- Impact Residual vs. Predicted
- Studentized Residuals
- Leverage Plot
- Quantile Plot (Q-Q Plot)
- Histogram of Residuals





RISKS

Over fitting

Data Quality

MITIGATION STRATEGIES

Cross Validation

Data Updates



Actionable Recommendations

- 🕒 Enhance Fuel Efficiency
- 🕒 Lightweight Materials
- 🕒 Stricter Emissions Standards
- 🕒 Incentives for Low-Emission Vehicles
- 🕒 Transparent Vehicle Labeling





THANK YOU

