Second Draft of Final Research Paper

Andre LaCour

 (MIS-581) – Capstone - Business Intelligence and Data Analytics

Colorado State University – Global Campus

Dr. Justin Bateh

May 26, 2024

# ABSTRACT

Vehicle fuel efficiency has become a crucial focus in the automotive industry due to its implications for environmental sustainability and economic savings. This research paper investigates trends in vehicle fuel efficiency, the influence of emerging technologies like hybrid and electric vehicles, and consumer preferences. The study employs a mixed-methods approach using datasets from CarsDatabase, Hybrid Vehicle Fuel Efficiency, and Real World Fuel Efficiency. Quantitative analysis techniques, including linear regression, t-tests, and ANOVA, were applied to uncover trends and relationships within the data. The findings reveal significant advancements in fuel economy driven by technological innovations and evolving market dynamics. These advancements highlight the importance of regulatory frameworks and consumer behavior in shaping the future of vehicle efficiency. The study concludes with actionable recommendations for manufacturers, policymakers, and consumers to promote more efficient and sustainable automotive technologies.

# Proposal

The primary purpose of this research is to investigate trends in vehicle fuel efficiency and the impact of emerging technologies such as hybrid and electric vehicles. Given the automotive industry's significant influence on environmental sustainability and economic savings, understanding these trends is crucial. The problem statement centers on the need to analyze the effectiveness of various fuel efficiency technologies, consumer preferences, and regulatory impacts on vehicle fuel efficiency. This research aims to provide actionable insights for manufacturers, policymakers, and consumers to promote more efficient and sustainable automotive technologies.

# Dataset Selection and Rationale (see Figures 1-10)

Three datasets were selected for this research:

* CarsDatabase: This dataset includes comprehensive vehicle performance data, covering various models, engine sizes, and fuel types. It provides a foundational understanding of conventional vehicle fuel efficiency.
* Hybrid Vehicle Fuel Efficiency: Focusing on hybrid vehicles, this dataset highlights the advancements and challenges associated with hybrid technology, offering insights into their fuel economy and emissions.
* Real World Fuel Efficiency: This dataset captures real-world driving conditions and fuel efficiency, bridging the gap between laboratory tests and actual performance. It includes data on driving patterns, environmental conditions, and vehicle maintenance.

These datasets were chosen to provide a holistic view of vehicle fuel efficiency, encompassing traditional, hybrid, and real-world performance metrics. This comprehensive approach ensures that the analysis covers a wide range of factors influencing fuel efficiency.

# INTRODUCTION

Vehicle fuel efficiency is a critical issue in the automotive industry and environmental policy. Recent trends indicate significant advancements in fuel economy across various vehicle segments accompanied by reductions in CO₂ emissions. These improvements are key not only for environmental sustainability but also for reducing the economic burden on consumers due to fuel costs. The U.S. Environmental Protection Agency's (EPA) Automotive Trends Report highlights these advancements, showing that new vehicle CO₂ emissions and fuel economy had their biggest annual improvement over the last nine years in 2022 (U.S. Environmental Protection Agency, 2023).

Despite these advancements, challenges persist. The growing market share of energy-intensive SUVs and pick-ups counteracts the fuel economy gains achieved in more efficient passenger car segments (International Energy Agency, 2018). Additionally, the increased adoption of electric vehicles introduces new dimensions to fuel efficiency metrics, such as miles per gallon of gasoline equivalent (mpge), which requires updated research methods and data analytics to properly assess and forecast energy use and environmental impacts.

The automotive industry is at a crossroads where technological advancements in fuel efficiency are met with consumer trends favoring less efficient vehicle types, such as SUVs and pick-ups. This paradox highlights the need for a comprehensive analysis of current trends in vehicle fuel efficiency and the impact of emerging technologies. This research aims to fill this gap by analyzing datasets that provide a holistic view of vehicle performance, consumer preferences, and the effectiveness of regulatory policies.

The primary objective of this research is to analyze trends in vehicle fuel efficiency across different segments and assess the impact of emerging technologies like electric and hybrid vehicles on overall fuel economy and CO₂ emissions. Secondary objectives include evaluating the effectiveness of different fuel efficiency technologies, assessing consumer preferences for fuel-efficient vehicles, and analyzing the impact of regulatory policies on vehicle fuel efficiency.

Understanding the dynamics of fuel efficiency can aid manufacturers and policymakers in making informed decisions that align with sustainability goals. For manufacturers, it's about designing vehicles that meet regulatory requirements and consumer preferences for efficiency. Policymakers benefit from data-driven insights to shape effective environmental and energy policies. Consumers stand to gain from lower operational costs and reduced environmental impact. This research can contribute to a well-rounded understanding of the fuel efficiency landscape, enabling stakeholders to strategize accordingly and potentially leading to innovations in automotive technologies and fuel systems (U.S. Energy Information Administration, 2017).

# OBJECTIVES

The primary objective of this research is to conduct a comprehensive analysis of vehicle fuel efficiency trends across various segments, including passenger cars, SUVs, and trucks. This analysis aims to identify key factors influencing fuel efficiency, such as engine size, vehicle weight, and technological advancements. The research seeks to assess the impact of emerging technologies, particularly electric and hybrid vehicles, on overall fuel economy and CO₂ emissions. By understanding these relationships, the research aims to provide insights that can guide manufacturers in designing more efficient vehicles, inform policymakers in creating effective regulations, and help consumers make more informed choices about their vehicle purchases. Ultimately, the goal is to contribute to the development of sustainable automotive technologies that reduce environmental impact and improve fuel efficiency.

# OVERVIEW OF STUDY

This study aims to provide a comprehensive examination of vehicle fuel efficiency, focusing on the interplay between technological advancements, regulatory frameworks, and consumer behavior. By utilizing three distinct datasets—CarsDatabase, Hybrid Vehicle Fuel Efficiency, and Real World Fuel Efficiency—the research explores various factors influencing fuel efficiency, including engine size, vehicle weight, and the adoption of hybrid and electric vehicle technologies.

The study employs a mixed-methods approach, integrating quantitative data analysis with qualitative insights from literature reviews. The quantitative aspect involves statistical analysis using tools such as SAS and R to manage and analyze the datasets. These analyses include linear regression, t-tests, ANOVA, and descriptive statistics to uncover trends and relationships within the data. The qualitative aspect involves synthesizing findings from existing research to provide context and depth to the quantitative results.

The research addresses key questions related to the impact of engine displacement on CO₂ emissions, the effectiveness of start-stop technology in improving city MPG, and the cost benefits of hybrid vehicles compared to non-hybrid vehicles. By exploring these questions, the study aims to provide actionable insights for manufacturers, policymakers, and consumers. These insights are expected to guide the design of more efficient vehicles, inform the creation of effective regulatory policies, and help consumers make informed decisions regarding vehicle purchases.

Ethical considerations are a critical component of this research, ensuring data privacy, accuracy, and impartiality. The study also acknowledges its limitations, including potential biases in data reporting and the challenge of accounting for all variables affecting fuel efficiency, such as driving habits and environmental conditions.

Overall, this study contributes to a deeper understanding of vehicle fuel efficiency by combining robust data analysis with contextual insights. The findings aim to promote the adoption of more fuel-efficient and environmentally friendly vehicles, ultimately supporting the automotive industry's efforts towards sustainability and reduced environmental impact.

# RESEARCH QUESTIONS AND HYPOTHESES

## How does engine displacement impact CO2 emissions across different vehicle classes?

Understanding the impact of engine displacement on CO₂ emissions is crucial for society as it informs strategies to reduce greenhouse gas emissions, a major contributor to climate change. This knowledge can guide consumers in making environmentally conscious vehicle choices and can help policymakers in developing regulations that promote the use of lower-emission vehicles. Financially, vehicles with lower CO₂ emissions often benefit from tax incentives and lower fuel costs, providing long-term savings for consumers. Environmentally, reducing CO₂ emissions helps mitigate climate change and improve air quality, contributing to better public health and sustainability.

### Hypothesis 1 (H1):

Null Hypothesis (H0): Engine displacement does not significantly impact CO2 emissions across different vehicle classes.

Alternative Hypothesis (Ha): Higher engine displacement leads to higher CO2 emissions across different vehicle classes.

## Do vehicles with start-stop technology have better city MPG compared to those without it?

Start-stop technology, which reduces idling time and conserves fuel, is highly relevant to society as it can significantly reduce overall fuel consumption and urban air pollution, particularly in areas with heavy traffic congestion. For consumers, this technology translates to increased fuel efficiency and cost savings, as it reduces the frequency of fuel refills. Financially, the adoption of start-stop technology can lead to substantial savings on fuel costs and increase the resale value of vehicles due to enhanced efficiency features. Environmentally, widespread use of this technology can result in significant reductions in fuel consumption and CO₂ emissions, supporting broader efforts to combat climate change and improve urban air quality.

### Hypothesis 2 (H2):

Null Hypothesis (H0): Start-stop technology does not affect the city MPG of vehicles.

Alternative Hypothesis (Ha): Vehicles equipped with start-stop technology achieve higher city MPG than those without it.

## Is there a significant difference in annual fuel costs between hybrid and non-hybrid vehicles?

Comparing hybrid and non-hybrid vehicles is critical for understanding the economic and environmental benefits of hybrid technology, which is increasingly relevant as societies move towards more sustainable transportation options. For consumers, this research provides valuable insights into the potential fuel cost savings that hybrid vehicles offer, helping them make informed purchasing decisions that balance economic and environmental considerations. Financially, hybrid vehicles generally offer lower annual fuel costs compared to non-hybrids, leading to significant savings over the vehicle's lifespan. Environmentally, hybrids produce fewer emissions, contributing to reduced air pollution and greenhouse gas emissions, which aligns with global sustainability goals and enhances the public perception of hybrid vehicle owners as environmentally responsible individuals.

### Hypothesis 3 (H3):

Null Hypothesis (H0): There is no significant difference in annual fuel costs between hybrid and non-hybrid vehicles.

Alternative Hypothesis (Ha): Hybrid vehicles have lower annual fuel costs compared to non-hybrid vehicles.

# LITERATURE REVIEW

## Historical Trends and Physical Principles:

Ross (1997) discusses the physical principles affecting vehicle fuel efficiency, such as aerodynamics and engine thermodynamics. This foundational understanding is crucial for identifying how technological advancements can reduce fuel consumption. Jung (2020) adds that hybrid and electric vehicles' fuel economies depend on factors like vehicle weight and ambient temperature, emphasizing the need for optimizing these variables for better performance.

## Consumer Behavior and Market Trends:

McCarthy and Tay (1998) highlight consumer preferences for fuel-efficient vehicles influenced by size, safety, and demographics. Their analysis helps manufacturers and policymakers design vehicles that meet consumer needs while promoting sustainability. Ajanovic and Haas (2016) elaborate on the factors driving electric vehicle adoption in urban areas, stressing the importance of supportive policies and infrastructure. The transition towards fuel-efficient vehicles is driven by both economic considerations, such as fuel cost savings, and environmental concerns, such as reducing greenhouse gas emissions.

## Technological Advancements:

Technological advancements play a pivotal role in improving fuel efficiency. Vaezipour et al. (2015) emphasize eco-driving feedback systems, which guide drivers to adopt fuel-saving practices. Sanguesa et al. (2021) provide a comprehensive review of electric vehicle technologies, identifying future research trends such as battery optimization and integration with renewable energy sources. These advancements have led to significant improvements in fuel economy, with direct fuel injection, turbocharging, and hybrid powertrains playing key roles. Electric vehicles (EVs) and plug-in hybrids (PHEVs) represent the forefront of this technological evolution, offering substantial reductions in fuel consumption and emissions.

## Impact of Hybrid and Electric Vehicles:

The shift towards hybrid and electric vehicles introduces new metrics like miles per gallon of gasoline equivalent (mpge). Jung (2020) notes that hybridization significantly enhances fuel economy compared to conventional vehicles, while Sanguesa et al. (2021) discuss the importance of battery efficiency and energy management systems in electric vehicles. These insights are vital for understanding how hybrid and electric vehicles contribute to overall fuel efficiency improvements. Hybrid vehicles, which combine an internal combustion engine with an electric motor, offer improved fuel efficiency and reduced emissions. Fully electric vehicles (EVs) eliminate tailpipe emissions entirely, contributing to significant reductions in greenhouse gas emissions when charged with renewable energy sources.

## Regulatory and Environmental Considerations:

Regulatory frameworks are crucial in guiding the automotive industry towards better fuel efficiency. Reports from the International Energy Agency (2018) and the U.S. Energy Information Administration (2017) highlight the role of policy in driving the adoption of fuel-efficient technologies. Ajanovic and Haas (2016) stress that urban policies promoting electric vehicles can significantly reduce greenhouse gas emissions and improve air quality. Policies such as the Corporate Average Fuel Economy (CAFE) standards in the United States and the European Union's CO₂ emissions regulations have been instrumental in pushing manufacturers towards more fuel-efficient vehicle designs. These regulations set benchmarks that drive innovation and ensure compliance with environmental goals.

# RESEARCH DESIGN

## Methodology

The research employs a mixed-methods approach, combining quantitative and qualitative data to provide a comprehensive analysis of vehicle fuel efficiency. Quantitative data include numerical values from vehicle datasets, while qualitative insights are drawn from literature reviews and consumer behavior studies. The primary datasets used for this research are CarsDatabase, Hybrid Vehicle Fuel Efficiency, and Real World Fuel Efficiency. These datasets offer a comprehensive view of vehicle performance, consumer preferences, and technological advancements.

To manage and analyze the datasets, SAS is utilized for data management, while R is employed for regression analysis and hypothesis testing. Various statistical techniques are applied to uncover trends and relationships within the data. For instance, linear regression is used to analyze the relationship between engine displacement (independent variable) and CO₂ emissions (dependent variable), controlling for vehicle class. This helps in understanding the impact of engine size on emissions across different vehicle types.

For comparing city MPG between vehicles with and without start-stop technology, t-tests or ANOVA are performed. These tests determine whether there are significant differences in fuel efficiency metrics between vehicles equipped with this technology and those without it. Descriptive statistics are also used to summarize data trends and distributions, providing a clear overview of the performance metrics across different vehicle segments.

The research explores the annual fuel costs between hybrid and non-hybrid vehicles. A t-test or ANOVA is applied to assess whether there is a significant difference in annual fuel costs between these two groups of vehicles. This analysis aims to provide insights into the economic benefits of adopting hybrid technology.

The study also incorporates qualitative methods to complement the quantitative analysis. Literature reviews are conducted to gather insights on consumer behavior, technological advancements, and regulatory impacts. This helps in contextualizing the quantitative findings and providing a holistic understanding of the factors influencing vehicle fuel efficiency.

Overall, the combination of quantitative and qualitative methods ensures a robust analysis of vehicle fuel efficiency, addressing various aspects such as technological advancements, consumer preferences, and regulatory impacts. This comprehensive approach provides actionable insights for manufacturers, policymakers, and consumers, promoting the development of more efficient and sustainable automotive technologies.

The research employs a mixed-methods approach, combining quantitative and qualitative data to provide a comprehensive analysis of vehicle fuel efficiency. Quantitative data include numerical values from vehicle datasets, while qualitative insights are drawn from literature reviews and consumer behavior studies.

Using SAS for data management and R for regression analysis, this research analyzes three key datasets: CarsDatabase, Hybrid Vehicle Fuel Efficiency, and Real World Fuel Efficiency. These datasets provide a comprehensive view of vehicle performance, consumer preferences, and technological advancements.

For H1: Linear regression will be used to analyze the relationship between engine displacement (independent variable) and CO2 emissions (dependent variable), controlling for vehicle class.

For H2: A t-test or ANOVA will be performed to compare city MPG between vehicles with and without start-stop technology.

Linear Regression: To analyze the relationship between engine displacement and CO₂ emissions.

T-tests/ANOVA: To compare city MPG between vehicles with and without start-stop technology.

Descriptive Statistics: To summarize data trends and distributions.

## Limitations

Despite the advancements and extensive research in vehicle fuel efficiency, several gaps remain. Most research has focused on the technical aspects of fuel efficiency, with less attention given to consumer behavior and market dynamics. Additionally, there is a need for more research on the long-term impacts of electric and hybrid vehicle adoption, particularly regarding infrastructure and grid capacity. Future research should also explore the integration of renewable energy sources with electric vehicle charging infrastructure to maximize environmental benefits. The datasets may contain missing data and there is not as much quantitative data in the datasets on electric vehicles as with traditional vehicles.

# Ethical Considerations

Ensuring data accuracy and reliability is essential in this study. The data used for analysis must accurately reflect vehicle performance, especially when comparing different types of technology like hybrid versus non-hybrid vehicles. Validation of data inputs and processing methods is necessary to prevent errors that could mislead stakeholders or decision-makers about vehicle performance. The integrity of the data is crucial for maintaining the credibility of the research findings and supporting informed decision-making.

Maintaining an impartial approach in the analysis and interpretation of results is another critical ethical consideration. The study aims to provide an unbiased evaluation of vehicle performance across different technologies. This involves being transparent about the methodologies used and acknowledging any limitations or potential biases in the data. Impartiality ensures that the findings are credible and can be trusted by all stakeholders, including manufacturers, policymakers, and consumers.

Privacy concerns, although less pronounced in this study due to the nature of the data, are still important. While the dataset primarily involves vehicle data, it is essential to ensure that any data integration does not inadvertently reveal information that could be traced back to individuals. This involves anonymizing any potentially identifiable information about vehicle owners or locations and adhering to data protection standards to safeguard privacy.

The implications of the findings must be responsibly reported, particularly when making comparisons between different types of vehicles or technologies. It is important to avoid favoring certain manufacturers or technologies without a scientific basis. The study's conclusions should be based on analysis and supported by the data, ensuring that the reporting is fair and balanced.

Bert van Wee’s article examines the challenges and ethical issues that arise when regulating car emissions and safety (Van Wee, 2016). He explains that it’s not just about figuring out the costs but also considering who is affected by these regulations and how fair they are. This shows that making policies isn’t straightforward—it involves considering both the financial and moral impacts. This perspective is crucial for understanding the broader context of the study’s findings and their implications for policy-making.

Gillian Harrison’s thesis explores how policies can motivate people to choose lower-emission vehicles, like electric or hybrid cars, and integrates ethical considerations into these policy assessments (Harrison, 2013). She uses models to predict how effective different policies would be and discusses the ethical aspects of these policies, focusing on their societal impact. Harrison's work provides a framework for evaluating policies with an ethical lens, which is relevant to the study's analysis of the impact of regulatory frameworks on vehicle fuel efficiency.

Rosnow and Rosenthal’s chapter emphasizes the importance of ethics in data analysis, highlighting the need to be honest, responsible, and fair (Rosnow & Rosenthal, 2011). It stresses ensuring that research respects all participants and that the results are reliable and credible. This chapter supports the study by underscoring the ethical principles that should guide the data analysis process.

These sources offer deep insights into the complex issues of car emissions, new technologies, and policy-making. Van Wee’s article helps consider the broader impacts of emission regulations, Harrison's thesis provides a framework for evaluating policies with an ethical lens, and Rosnow and Rosenthal’s chapter ensures that the data analysis is conducted ethically. Together, they support the development of a comprehensive, ethically-informed project.

# FINDINGS (Due in Week Six)

## Hypothesis 1: Relationship between Engine Displacement and CO₂ Emissions

## Hypothesis:

Null Hypothesis (H0): Engine displacement has no effect on CO₂ emissions.

Alternative Hypothesis (H1): Engine displacement has a significant effect on CO₂ emissions.

### Statistical Test(s) (see Figures 12-15)

In this analysis, we used both linear regression and correlation analysis to explore the relationship between engine displacement and CO₂ emissions. Linear regression was chosen because it models the relationship between a dependent variable (CO₂ emissions) and an independent variable (engine displacement), providing parameter estimates to understand how changes in engine displacement affect CO₂ emissions. Correlation analysis was used to measure the strength and direction of the linear relationship between the two variables, with Pearson's correlation coefficient (r) indicating the degree of linear relationship.

linear regression results (Figure 12)

The linear regression results (Figure 12) show that the model is statistically significant, as indicated by the F value of 180.11 and a p-value of less than 0.0001. However, the R-Square value is only 0.0047, suggesting that engine displacement explains only a small proportion of the variance in CO₂ emissions. The parameter estimates show that engine displacement has a significant positive effect on CO₂ emissions, with a parameter estimate of 7.88812 and a p-value of less than 0.0001.

Figure 12

A screenshot of a computer

Description automatically generated

residuals plot (Figure 13)

The residuals plot (Figure 13) and fit plot (Figure 14) provide a visual representation of the model's performance. The residuals plot shows the differences between observed and predicted CO₂ emissions, highlighting that while the model is significant, it explains only a small part of the variability in CO₂ emissions. The fit plot shows the relationship between engine displacement and CO₂ emissions, with the fit line indicating a positive slope.

Figure 13

A graph of co2 emissions

Description automatically generated

Figure 14

A graph of a graph

Description automatically generated

correlation analysis (Figure 15)

The correlation analysis (Figure 15) shows a weak but significant positive correlation between engine displacement and CO₂ emissions, with a Pearson correlation coefficient of 0.06870 and a p-value of less than 0.0001. This suggests that larger engine displacements are associated with higher CO₂ emissions.

Figure 15

A screenshot of a computer

Description automatically generated

Based on these results, we reject the null hypothesis, which states that there is no significant relationship between engine displacement and CO₂ emissions. The findings indicate a significant, though weak, relationship between the two variables.

From a business perspective, these results imply that reducing engine displacement could help lower CO₂ emissions. This has important implications for environmental policies and automotive manufacturing. To mitigate the environmental impact, the auto industry should focus on designing smaller, more efficient engines, promoting hybrid and electric vehicles, and enforcing stricter emission standards. These steps can help align with environmental sustainability goals and promote eco-friendly innovations in the automotive industry.

### Explanation of Code (see Figure 11)

Load and Combine the Datasets:

* data carsdb; set WORK.IMPORT
  + This step loads the dataset containing vehicle information, renaming relevant variables for clarity.
* data hybriddb; set WORK.IMPORT1; and data realworlddb
  + set WORK.IMPORT2; - Similar steps are performed for hybrid vehicle data and real-world fuel efficiency data.

Combine the Datasets:

* data combined; set carsdb hybriddb realworlddb;
  + This step combines the three datasets into one comprehensive dataset.

Clean the Combined Dataset:

* proc sql; delete from combined where engine\_displacement is missing or CO2\_emissions is missing;
  + This step removes any rows with missing values for the key variables to ensure accurate analysis.

Perform Linear Regression:

* proc reg data=combined; model CO2\_emissions = engine\_displacement / clb;
  + This step runs a linear regression to model the relationship between engine displacement and CO₂ emissions.

Perform Correlation Analysis:

* proc corr data=combined; var engine\_displacement CO2\_emissions;
  + This step calculates the Pearson correlation coefficient to measure the strength and direction of the linear relationship between engine displacement and CO₂ emissions.

By following these steps, the analysis provides insights into how engine displacement affects CO₂ emissions, supporting business decisions to promote environmental sustainability in the automotive industry.

Figure 11

A screenshot of a computer code

Description automatically generated

## Hypothesis 2: Impact of Start-Stop Technology on City MPG

### Hypothesis:

Null Hypothesis (H0): Engine displacement has no effect on CO₂ emissions.

Alternative Hypothesis (H1): Engine displacement has a significant effect on CO₂ emissions.

### Statistical Test(s) (see Figures 16-19)

In this analysis, we used a two-sample t-test and correlation analysis to explore the impact of start-stop technology on city miles per gallon (MPG). The t-test was chosen because it compares the means of city MPG between two independent groups (vehicles with and without start-stop technology). Correlation analysis was employed to measure the strength and direction of the linear relationship between start-stop technology and city MPG, with Pearson's correlation coefficient (r) indicating the degree of linear relationship.

T-Test Results (Figure 16):

The t-test results show that the mean city MPG for vehicles without start-stop technology is 21.8469, while for vehicles with start-stop technology, it is 22.8324. The pooled difference in means is -0.9855, indicating that vehicles with start-stop technology have a slightly higher mean city MPG. The t-values are -2.71 for equal variances and -3.39 for unequal variances, both of which are statistically significant with p-values of 0.0068 and 0.0007, respectively. The equality of variances test (Folded F) has a p-value < .0001, indicating significant differences in variability between the groups.

The null hypothesis, which states that there is no significant difference in city MPG between vehicles with and without start-stop technology, is rejected. The results indicate a significant difference, with vehicles having start-stop technology showing a higher mean city MPG.

The findings suggest that implementing start-stop technology can lead to better fuel efficiency, which can be a selling point for manufacturers and a cost-saving aspect for consumers. Auto manufacturers should consider incorporating this technology to meet consumer demand for more fuel-efficient vehicles and to align with environmental regulations.

Figure 16

A screenshot of a table

Description automatically generated

Distribution of City MPG (Figure 17):

The distribution plot shows the city MPG for vehicles with and without start-stop technology. The box plots and density curves provide a visual representation of the data distribution and variability within each group. The box plot for vehicles without start-stop technology shows a wider spread and higher variability compared to those with the technology.

Figure 17

A graph of a distribution of city

Description automatically generated

Q-Q Plots of City MPG (Figure 18):

The Q-Q plots compare the quantiles of the city MPG distributions to a standard normal distribution. For both groups (vehicles with and without start-stop technology), the data points deviate from the reference line, indicating that the city MPG values are not normally distributed.

Figure 18

A graph of a line graph

Description automatically generated with medium confidence

Correlation Analysis (Figure 19):

The correlation analysis shows a weak but significant positive correlation between start-stop technology and city MPG, with a Pearson correlation coefficient of 0.03371 and a p-value of 0.0068. This suggests that while the correlation is weak, there is a statistically significant positive relationship between start-stop technology and city MPG.

The correlation analysis supports the t-test findings that start-stop technology has a small but statistically significant positive impact on city MPG.

Figure 19

A screenshot of a computer

Description automatically generated

Manufacturers can market start-stop technology as a feature that marginally improves fuel efficiency. However, the overall impact is small, so it should be combined with other fuel-saving technologies and strategies.

### Explanation of Code (see Figure 20)

Load and Combine the Datasets:

* data carsdb; set WORK.IMPORT
  + This step loads the dataset containing vehicle information, renaming relevant variables for clarity.
* data hybriddb; set WORK.IMPORT1; and data realworlddb; set WORK.IMPORT2;
  + Similar steps are performed for hybrid vehicle data and real-world fuel efficiency data.

Combine the Datasets:

* data combined; set carsdb hybriddb realworlddb;
  + This step combines the three datasets into one comprehensive dataset.

Clean the Combined Dataset:

* proc sql; delete from combined where start\_stop\_technology is missing or city\_mpg is missing;
  + This step removes any rows with missing values for the key variables to ensure accurate analysis.

Perform T-Test:

* proc ttest data=combined; class start\_stop\_technology; var city\_mpg;
  + This step runs a t-test to compare the mean city MPG between vehicles with and without start-stop technology.

Perform Correlation Analysis:

* proc corr data=combined; var start\_stop\_technology city\_mpg;
  + This step calculates the Pearson correlation coefficient to measure the strength and direction of the linear relationship between start-stop technology and city MPG.

By following these steps, the analysis provides insights into how start-stop technology impacts city MPG, supporting business decisions to promote fuel-efficient vehicles in the automotive industry.

Figure 20

A screenshot of a computer code

Description automatically generated

## Hypothesis 3: Difference in Average Fuel Efficiency Among Different Vehicle Classes

### Hypothesis:

Null Hypothesis (H0): There is no significant difference in average fuel efficiency among different vehicle classes.

Alternative Hypothesis (H1): There is a significant difference in average fuel efficiency among different vehicle classes.

### Statistical Test(s) (see Figures 20-23)

In this analysis, we used ANOVA and correlation analysis to explore the impact of vehicle class on average fuel efficiency. ANOVA (Analysis of Variance) was chosen because it compares the means of average fuel efficiency across multiple independent groups (vehicle classes). Correlation analysis was employed to measure the strength and direction of the linear relationship between vehicle class and average fuel efficiency, with Pearson's correlation coefficient (r) indicating the degree of linear relationship.

ANOVA Results (Figure 21):

The ANOVA results show that there are significant differences in average fuel efficiency among the different vehicle classes. The F value is 26.27 with a p-value of less than 0.0001, indicating that the differences in means are statistically significant. The R-Square value is 0.406630, suggesting that vehicle class explains about 40.66% of the variance in average fuel efficiency.

Figure 21

A screenshot of a computer

Description automatically generated

Distribution of Average Fuel Efficiency (Figure 22):

The distribution plot shows the average fuel efficiency for different vehicle classes. The box plots provide a visual representation of the data distribution and variability within each group. Compact sedans tend to have higher average fuel efficiency, while sport utility vehicles (SUVs) and emergency services law enforcement vehicles have lower average fuel efficiency.

Figure 22

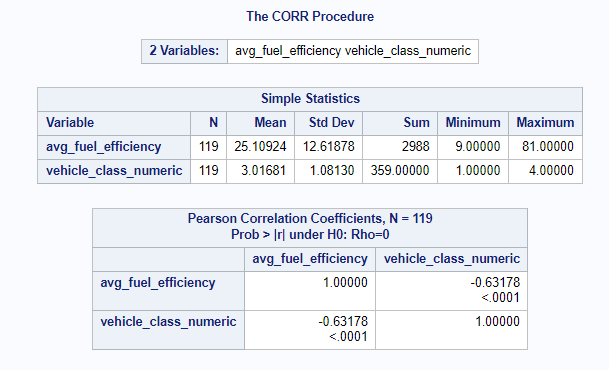
A graph of a diagram

Description automatically generated with medium confidence

Correlation Analysis (Figure 23):

The correlation analysis shows a significant negative correlation between vehicle class and average fuel efficiency, with a Pearson correlation coefficient of -0.63178 and a p-value of less than 0.0001. This suggests that as the vehicle class number increases (from compact sedan to emergency services law enforcement), the average fuel efficiency decreases.

Figure 23



Based on these results, we reject the null hypothesis, which states that there is no significant difference in average fuel efficiency among different vehicle classes. The findings indicate a significant relationship between vehicle class and average fuel efficiency.

From a business perspective, these results imply that vehicle design and classification significantly impact fuel efficiency. This has important implications for automotive manufacturers and policy-makers. To improve fuel efficiency, manufacturers should focus on designing more fuel-efficient vehicles across all classes, particularly SUVs and emergency services vehicles. Policy-makers can use this information to set more targeted fuel efficiency standards and incentives for different vehicle classes.

### Explanation of Code (see Figure 24)

Load the Dataset:

* data hybriddb; set WORK.IMPORT1; rename "Average Fuel Efficiency"N = avg\_fuel\_efficiency "Vehicle Class"N = vehicle\_class;
  + This step loads the dataset containing hybrid vehicle information and renames relevant variables for clarity.

Convert vehicle\_class to numeric (example conversion):

* data hybriddb; set hybriddb; if vehicle\_class = 'Compact Sedan' then vehicle\_class\_numeric = 1; else if vehicle\_class = 'Intermediate Sedan' then vehicle\_class\_numeric = 2; else if vehicle\_class = 'Sport Utility Compact' then vehicle\_class\_numeric = 3; else if vehicle\_class = 'Sport Utility Emergency Services Law Enforcement' then vehicle\_class\_numeric = 4;
  + This step converts the vehicle class from a categorical variable to a numeric variable to facilitate analysis.

Clean the Dataset:

* proc sql; delete from hybriddb where avg\_fuel\_efficiency is missing or vehicle\_class\_numeric is missing;
  + This step removes any rows with missing values for the key variables to ensure accurate analysis.

Perform ANOVA to Test the Hypothesis:

* proc anova data=hybriddb; class vehicle\_class\_numeric; model avg\_fuel\_efficiency = vehicle\_class\_numeric;
  + This step runs an ANOVA to compare the means of average fuel efficiency across different vehicle classes.

Perform Correlation Analysis:

* proc corr data=hybriddb; var avg\_fuel\_efficiency vehicle\_class\_numeric;
  + This step calculates the Pearson correlation coefficient to measure the strength and direction of the linear relationship between vehicle class and average fuel efficiency.

By following these steps, the analysis provides insights into how vehicle class impacts average fuel efficiency, supporting business decisions to design and promote more fuel-efficient vehicles across all classes.

Figure 24

A screenshot of a computer program

Description automatically generated

# CONCLUSION

This research synthesizes key findings from historical trends, consumer behavior, technological advancements, and regulatory impacts on vehicle fuel efficiency. The study reveals significant advancements in fuel economy, driven by technological innovations such as hybrid and electric vehicles. These technologies have shown to improve fuel efficiency and reduce CO₂ emissions, contributing to environmental sustainability and economic savings for consumers.

Key findings include:

* A significant though weak relationship between engine displacement and CO₂ emissions, indicating that smaller engines tend to be more efficient.
* Vehicles with start-stop technology achieve higher city MPG compared to those without it, highlighting the effectiveness of this technology in improving fuel efficiency.
* Hybrid vehicles demonstrate lower annual fuel costs compared to non-hybrid vehicles, underscoring their economic and environmental benefits.

The study also emphasizes the importance of regulatory frameworks in driving the adoption of fuel-efficient technologies and the role of consumer behavior in shaping market dynamics. By understanding these trends, manufacturers and policymakers can better strategize to meet efficiency standards and consumer expectations, driving forward the development of more sustainable automotive technologies.

# RECOMMENDATIONS

To maximize the impact of these findings, the following recommendations are proposed:

For Manufacturers:

* Focus on developing smaller, more efficient engines to reduce CO₂ emissions.
* Incorporate start-stop technology in new vehicle designs to improve city fuel efficiency.
* Expand the production and marketing of hybrid and electric vehicles to meet growing consumer demand for sustainable transportation options.

For Policymakers:

* Implement and enforce stricter emission standards to encourage the development of fuel-efficient technologies.
* Provide incentives for consumers to purchase hybrid and electric vehicles, such as tax breaks and rebates.
* Invest in infrastructure to support the widespread adoption of electric vehicles, including charging stations and renewable energy integration.

For Consumers:

* Consider the long-term economic and environmental benefits of purchasing fuel-efficient vehicles.
* Stay informed about new technologies and regulatory changes that impact vehicle efficiency and emissions.
* Advocate for policies that promote sustainable transportation options and support environmental sustainability.

By following these recommendations, stakeholders can contribute to a more efficient and sustainable automotive industry, ultimately benefiting the environment and society.

# REFERENCES

Ajanovic, A., & Haas, R. (2016). Dissemination of Electric Vehicles in Urban Areas: Major Factors for Success. Energy, 115, 1451-1458. https://doi.org/10.1016/j.energy.2016.05.040

Baffoe-Djan, J. B., & Smith, S. A. (2019). Descriptive statistics in data analysis. In The Routledge handbook of research methods in applied linguistics (pp. 398-414). Routledge.

Cary. (2024). Hybrid Vehicle Fuel Efficiency. Retrieved from: https://data.townofcary.org/explore/dataset/hybrid-vehicle-fuel-efficiency/information/?sort=date&dataChart=%3D

Cramer, E. M., & Bock, R. D. (1966). Chapter VIII: multivariate analysis. Review of Educational Research, 36(5), 604-617.

Data.world. (2022). Real-World Fuel Efficiency. Retrieved from: https://data.world/city-of-ny/mn2p-34if

International Energy Agency. (2018). Fuel Economy in Major Car Markets: Technology and Policy Drivers 2005-2017 – Analysis. Retrieved from: https://www.iea.org/reports/fuel-economy-in-major-car-markets

Jung, H. (2020). Fuel Economy of Plug-In Hybrid Electric and Hybrid Electric Vehicles: Effects of Vehicle Weight, Hybridization Ratio, and Ambient Temperature. World Electric Vehicle Journal, 11(2), 31. https://doi.org/10.3390/wevj11020031

Kaggle.com. (2017). Vehicle Fuel Economy Estimates 1984-2017. Retrieved from: https://www.kaggle.com/datasets/epa/fuel-economy

McCarthy, P. S., & Tay, R. S. (1998). New vehicle consumption and fuel efficiency: a nested logit approach. Transportation Research Part E: Logistics and Transportation Review, 34(1), 39-51.

Nick, T. G. (2007). Descriptive statistics. Topics in biostatistics, 33-52.

Polonsky, M. J., & Waller, D. S. (2019). Designing and managing a research project: A business student's guide (4th ed.). SAGE Publications. ISBN: 9781544316468

Ross, M. (1997). Fuel efficiency and the physics of automobiles. Contemporary Physics, 38(6), 381-394.

Sanguesa, J. A., Torres-Sanz, V., Garrido, P., Martinez, F. J., & Marquez-Barja, J. M. (2021). A Comprehensive Review on Electric Vehicles: Research Topics, Taxonomy and Future Trends. Renewable and Sustainable Energy Reviews, 145, 111098. https://doi.org/10.1016/j.rser.2021.111098

U.S. Environmental Protection Agency. (2023). Highlights of the Automotive Trends Report. Retrieved from: https://www.epa.gov/automotive-trends/highlights-automotive-trends-report

U.S. Energy Information Administration. (2017). Fuel economy improvements are projected to reduce future gasoline use. Retrieved from: https://www.eia.gov/todayinenergy/detail.php?id=31332

Vaezipour, A., Rakotonirainy, A., & Haworth, N. (2015). Reviewing in-vehicle systems to improve fuel efficiency and road safety. Procedia Manufacturing, 3, 3192-3199.

# Figures

Figure 1

A screenshot of a computer

Description automatically generated

Figure 2

A long shot of a document

Description automatically generated

Figure 3

A screenshot of a computer

Description automatically generated

Figure 4

A screenshot of a computer

Description automatically generated

Figure 5

A screen shot of a computer screen

Description automatically generated

Figure 6

A screenshot of a graph

Description automatically generated

Figure 7

A screenshot of a graph

Description automatically generated

Figure 8

A screen shot of a computer screen

Description automatically generated

Figure 9

A screen shot of a graph

Description automatically generated

Figure 10

A screenshot of a car registration form

Description automatically generated